REPORT RESUMES

ED 020 683 EM 006 688

LINEAR STRUCTURAL MODELS FOR RESPONSE AND LATENCY PERFORMANCE IN ARITHMETIC. PSYCHOLOGY SERIES, TECHNICAL REFORT NO. 100.

BY- SUPPES, PATRICK AND OTHERS
STANFORD UNIV., CALIF. INST. FOR MATH. STUDIES

PUB DATE 29 JUL 66

EDRS PRICE MF-$0.50 HC-$4.12 101P.

DESCRIPTORS- *LEARNING THEORIES, PREDICTIVE ABILITY (TESTING), *PERFORMANCE FACTORS, ELEMENTARY SCHOOL MATHEMATICS, *RESPONSE MODE, LEARNING PROCESSES, COGNITIVE PROCESSES, *ARITHMETIC, COMPUTER ASSISTED INSTRUCTION, TIME SHARING, CONCEPT FORMATION, TEACHING MACHINES, PROGRAMED INSTRUCTION, GRADE 3, GRADE 4, GRADE 5, GRADE 6, *PROBLEM SOLVING, MEMORY, FACTORY ANALYSIS, MODELS,

A LEARNING MODEL TO IDENTIFY FACTORS CONTRIBUTING TO THE DIFFICULTY OF A PROBLEM ITEM WAS SUPPORTED EMPIRICALLY, AND INDICATED THAT THE NUMBER OF STEPS REQUIRED TO SOLVE A PROBLEM WAS THE MOST IMPORTANT VARIABLE IN PREDICTING BOTH ERROR PROBABILITY AND RESPONSE LATENCY. THE MODEL, IN ORDER TO ESTABLISH DIFFERENTIAL PREDICTIONS OF DIFFICULTY IN SOLVING ARITHMETIC PROBLEMS, IDENTIFIED SUCH VARIABLES AS THE MAGNITUDE OF THE LARGEST AND THE SMALLEST NUMBERS APPEARING, THE FORM OF THE EQUATION IN WHICH THE PROBLEMS ARE PRESENTED, AND THE DIFFERENCES BETWEEN NUMBERS IN SUBSTRACTIONS. ANOTHER VARIABLE, NUMBER OF STEPS (NSTEPS) REQUIRED TO SOLVE THE PROBLEMS, WAS FURTHER DIVIDED INTO TRANSFORMATION (STEPS REQUIRED TO PUT THE EQUATION INTO CANONICAL FORM), OPERATION (NUMBER OF OPERATIONS PERFORMED), AND MEMORY (NUMBER OF DIGITS THAT MUST BE HELD IN MEMORY). TERMINALS WERE INSTALLED IN EIGHT CLASSROOMS, AND 270 THIRD, FOURTH, FIFTH, AND SIXTH GRADERS PARTICIPATED FOR ONE ACADEMIC YEAR IN COMPUTER-ASSISTED MATHEMATICS INSTRUCTION. A REGRESSION ANALYSIS SHOWED NSTEPS TO BE THE MOST IMPORTANT VARIABLE IN PREDICTING BOTH ERROR PROBABILITY AND RESPONSE LATENCY, AND ANALYSIS OF THE FACTORS IN NSTEPS SHOWED THAT MEMORY WAS THE MOST IMPORTANT FACTOR. OPERATION PLAYED NO ROLE IN PREDICTING THE DEPENDENT VARIABLES. (OH)

ED020683

LINEAR STRUCTURAL MODELS FOR RESPONSE AND LATENCY PERFORMANCE IN ARITHMETIC

by

Patrick Suppes, Lester Hyman, and Max Jerman

TECHNICAL REPORT NO. 100

July 29, 1966

PSYCHOLOGY SERIES

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

EM 006 6

LINEAR STRUCTURAL MODELS FOR RESPONSE AND LATENCY PERFORMANCE IN ARITHMETIC*

Patrick Suppes, Lester Hyman, Max Jerman

Institute for Mathematical Studies in the Social Sciences

Stanford University, Stanford, California

1. Introduction.

In the cognitive domain mathematics provides one of the clearest examples of complex learning and performance, for the structure of the subject itself provides numerous constraints on any adequate theory. The learning and performance models derived from the main trends of contemporary mathematical learning theory have provided an excellent predictive analysis of a large variety of experimental situations. Unfortunately, however, most of these experimental situations are much simpler in structure than what corresponds to even the simplest parts of elementary mathematics. Because this claim is central to the motivations behind the present paper, we would like to expand on it in some detail.

---

*The research reported here has been supported by the U.S. Office of Education, the National Science Foundation and the Carnegie Corporation of New York. We are indebted to Guy Groen for a number of useful comments, to Manuel Uy for extensive and much appreciated efforts in executing the computer programs used in the data analysis; and above all to Dow Brian and his staff of computer programmers, including particularly John Gwinn, in the Institute's Computer-Based Laboratory for Learning and Teaching. Experiments of the magnitude reported here would have been a practical impossibility without the hardware and software facilities of the Laboratory. We are also indebted to Miss Luanne D. Berkowitz who assumed major responsibilities in running the experiments and in helping construct the curriculum materials, and to Mr. Victor Norton, the Principal of Grant Elementary School of Cupertino Union School District in California, as well as to the teachers of Grant, all of whom contributed substantially to the rather hectic beginning stages of initiating a new technology in an elementary-school setting.

The familiar and now classical linear model provides a good starting point for our discussion. For our purposes, we may take this model in its simplest form, as applied to a situation in which a given response is always reinforced, and all other responses are indicated as incorrect. For the formulation under this restriction let q_{n+1} be the probability of an incorrect response on trial $n + 1$. This probability is then the following simple linear function of the probability q_n of an incorrect response on trial n:

$$q_{n+1} = \alpha q_n ,$$

where the learning parameter α is such that $0 \leq \alpha < 1$. The formal properties of this simple model have now been investigated thoroughly and are well understood. It is apparent, however, that if the subject must learn a number of different items which differ in structure and therefore in learning difficulty, the simple linear model can accommodate this fact only by separately estimating a learning parameter α for each item. From the standpoint of classical psychological concerns with the character of learning and performance, this is far from satisfactory. What is desired, rather, is an analysis of the factors in the structure of the stimulus item which lead to varying difficulty. The estimation of a nonstructural parameter unique for each item is a way of handling data when no better resources are available, but it does not take us very deeply into the psychological problems of learning complex items like those common in mathematics and other structured subjects. Above all, the estimation of a separate parameter for each item leads to a wasteful use of parameters. In general, if we take a collection of items from a given domain of mathematics, we would like to be able to attach weights to the various factors that may be objectively identified in the item, and then use estimates of a few such weights to predict the relative difficulty or the latency of response for a large number of items. The linear model itself cannot provide such mechanisms. This is not to denigrate the importance and significance of the linear model, for it will doubtless enter in many places to provide an analysis of particular mechanisms. But it will not serve as anything like the basis for a fundamental or general theory of complex learning.

At first glance, it might appear that we could use a learning theory with more structure, such as stimulus-sampling theory, to provide an adequate analysis of stimulus structure--adequate to make differential predictions of difficulty in cognitive domains like that of elementary mathematics. An examination of the explicit axiomatizations of stimulus-sampling theory, which may be found in Estes and Suppes (1959), Suppes and Atkinson (1960) or Atkinson and Estes (1963), shows, however, that the concept of stimulus used does not provide an adequate analysis of structure. Roughly speaking, the situation is the following. The stimuli presented to a subject on a given trial are represented by a set of stimulus elements. In the concept of an arbitrary set of stimulus elements, there is the beginning of an adequate apparatus for the concept of structure, but the additional assumptions needed for a definite notion of structure have not yet been imposed on the concept of an arbitrary set. It is necessary to go beyond the current formulations of the theory in order to analyze even the simplest sort of stimulus items used in the teaching of elementary mathematics. Probably the most successful version of stimulus-sampling theory for a wide variety of experiments is the pattern conception of stimulus conditioning that originates with Estes (1959). On this theory, the individual stimulus elements are not conditioned as components to a correct response, rather, an entire pattern of stimulus elements is so conditioned, and in general the number of patterns available for sampling in a given stimulus situation will be a parameter to be estimated from the data. But even these conceptions are very far from providing an analysis sufficiently structured to yield differential predictions of difficulty in responding correctly to problem-items drawn from concepts and topics in elementary mathematics.

It might also be thought that the applications of stimulus-sampling theory or related sorts of theories to stimulus-discrimination problems during the past decade would yield theoretical ideas adequate to the analysis of complex structure. Again, however, an examination of the kinds of problems that have been handled shows very quickly that a structural apparatus adequate to problems in simple addition, for example, is certainly not even implicitly inherent in the theories that have been developed within the general framework of theories of conditioning.

Both psychologists and educators interested in cognitive theories in learning would undoubtedly very much agree with the remarks we have just made about stimulus-response theories. However, we find that we must say the same sorts of things about the current cognitive theories of learning and performance, which have attracted considerable interest in the last few years. As opposed to the stimulus-response theories that we have mentioned, perhaps the greatest defect of the cognitive theories is simply a lack of sufficient intellectual definiteness even to settle the question of whether or not specific predictions can be made. The kinds of cognitive considerations, for example, that enter into the studies reported in the well-known book by Bruner, Goodnow and Austin (1956) simply do not provide a framework within which we can ask specific questions about the estimation of parameters for the prediction of differential difficulty over a selection of stimulus items drawn from some complex domain, whether it be elementary mathematics or elementary language learning. Again we would not want to be misunderstood on this point. The analysis of the types of strategies used in concept attainment is certainly a useful contribution to the psychology of concept formation and thinking, but it must be realistically asserted that no theory has yet been sufficiently developed to provide the kind of parametric predictions that are considered a minimum requirement in the area of mathematical models of learning and performance. The same sorts of remarks apply to the invaluable work of Piaget and his collaborators. Piaget has contributed much to our understanding of cognitive development in children and especially to our understanding of the kind of structures children find or if you wish, create, in the stimulus environment. But again, Piaget's concepts have not been sufficiently articulated into a well-defined theory to provide parametric predictions of differential difficulty for items drawn from any cognitive domain. This is not particularly Piaget's task, as it was not Bruner's. Nevertheless, we do intend our remarks to be of a critical nature, for until parametric predictions can be produced from the theoretical proposals generated by various psychologists, these theoretical ideas cannot be accepted as a final analysis of what we hope to understand about cognitive processes.

The preceding remarks have mainly emphasized the inadequacy of current psychological theories to provide parametric predictions of differential difficulty as measured by the rate of correct responding. These theories are even more inadequate when we turn to response latencies. From the standpoint of the analysis of performance, latencies are in many respects more important as a source of information to the theorist than response data. This is particularly true of any studies devoted to skill performance after a good deal of learning has taken place. As some of the data reported here show, and as one would expect anyway on a priori grounds, the range of latencies observed in a group shows systematic variation in a way that clearly reflects a measure of item difficulty. What is ultimately desired in this case is the kind of model that can predict from the structure of an item the process a subject must go through in finding the correct response. In the case of arithmetic, at least part of this process must undoubtedly be related to the standard algorithms taught as part of the curriculum; but even a casual glance at these algorithms will show that the conception of them used in teaching and in the curriculum does not provide a sufficient analysis of processing to make differential predictions of difficulty as reflected in response latencies.

What is also surprising about latency is that there have been so few studies that reported detailed data on this measure. The only directly relevant studies that we have found in the literature on arithmetic are Batson and Combellick (1925), Helseth (1927), Knight and Behrens (1928) and Billington (1947). This absence of latency studies (even though there are undoubtedly several of which we are not aware) indicates how superficial has been the investigation of structural models adequate to predict differential difficulty either in terms of responses or response latencies.

The constructive aim of the present paper is to formulate and test some linear structural models that do lead to parametric predictions of the sort we have been discussing. The sense in which these models are linear is not precisely the same sense that applies to the linear learning model; but it is in the context of linear-regression models, a point that is made clear in the next section. The models and accompanying theory which we present and test in this paper are meant only as a beginning. We do believe that they provide a significant and promising foundation for further work.

2. The Theory.

The learning models that arise in stimulus-sampling theory all exemplify a certain class of stochastic processes, and in general a different class of such processes is exemplified by the linear models discussed at the beginning. In the same fashion, the linear structural models proposed in this paper all exemplify a general class of models that are classical in statistics. But simply to say that we are applying linear-regression models to the study of arithmetic performance provides no more clue to the theoretical ideas behind the analysis than does the assertion that we apply to a given body of learning phenomena a finite state Markov chain as the primary mathematical tool of analysis. What is important and significant for psychology is the particular way in which the broad class of linear-regression models is narrowed and made meaningful from the standpoint of response or latency performance in arithmetic.

It will perhaps be useful to begin with a class of problems that are simpler than those considered here in detail. The discussion of this first example follows Suppes (1966). Let us suppose that a set of problems consists only of simple addition problems of the following sort: $1 + 2 = n$, $1 + n = 3$ and $n + 2 = 3$. Let us restrict the sums to those not greater than 5. We postulate that the following five facts are held in memory:

1 = /

2 = //

3 = ///

4 = ////

5 = ///// .

Our algorithm is then the following:

(1) Replace all Arabic numerals by their stroke definitions and delete all plus symbols.

(2) If there are strokes on both sides of the equal sign, cancel one by one, starting from the left of each side until there remain no strokes on one side. Ignore n in cancelling.

(3) On the one side still having strokes, replace the strokes by an Arabic numeral, using the definitions in memory.

The solution in the form $n = c$ or $c = n$ will result.

To obtain a single factor f representing the number of steps, we simply count the number of steps required by the algorithm to solve a given problem. For example, the steps to solve $3 + n = 5$ are 5 in number.

(1) /// n = ///// by rule (1)

(2) // n = //// by rule (2)

(3) / n = /// by rule (2)

(4) n = // by rule (2)

(5) n = 2 by rule (3) ,

and thus for this model and this problem, $f = 5$. A more realistic version of this algorithm, at least for many standard situations in which students are tested on their command of the simple addition facts, is to postulate that the student counts the difference n, by beginning at 3 and stopping at 5. A test of five variants of this latter counting algorithm is reported in Suppes and Groen (1966).

For the problem-items analyzed in this paper the central problem is to identify the factors that contribute to the difficulty of the item. Typical factors that we shall examine are the magnitude of the largest number appearing in the problem, the magnitude of the smallest number, the form of the equation in which the problem is presented, and most importantly, the number of steps required to solve the problem. Exactly how the number of steps is to be defined is a matter that we take up in detail below. As a matter of notation we shall denote the j^{th} factor of problem i in a given set of problems or exercises by f_{ij}. The statistical parameters that must be estimated from the data are the weights to be attached to each factor. We shall denote the weight assigned to the j^{th} factor by α_j. We want to emphasize as explicitly as possible that the factors identified and used in the models presented in this paper are never factors in the sense of factor analysis; that is, the factors do not arise as abstract constructions from the data. Rather, they are always objective factors identifiable by the experimenter in the problem-items themselves, independent of any data analysis. Which factors turn out to be important is a matter of the estimated weights α_j, but in no case does the decision as to what is the numerical value of a factor for a given problem-item depend in any way on the response data themselves. In fact, it will be apparent that all of the factors used in the analyses presented here have an intuitive and direct relevance to commonsense ideas of difficulty, and their definitions are so straightforward and simple that there is little prospect of disagreement over their objective value in a given problem-item.

We may first consider the analysis of response data. Let p_i be the observed proportion of correct responses on problem-item i for a given group of subjects. The central task of a model is to predict the observed proportions p_i. The natural linear-regression model in terms of the factors f_{ij} and the weights α_j is simply

$$p_i = \sum_j \alpha_j f_{ij} + \alpha_o .$$

However, there is a central difficulty with this particular model: there is no guarantee that probability will be preserved as the estimated weightings and identifiable factors are combined to predict the observed proportion of correct responses in new items. Consequently, in order to guarantee preservation of probability, that is, to ensure that the predicted p_i's will always lie between 0 and 1, it is natural to make the following transformation and to define a new variable z_i,[1]

$$z_i = \log \frac{1 - p_i}{p_i} . \qquad (1)$$

And then to use as the regression model

$$z_i = \sum_j \alpha_j f_{ij} + \alpha_o . \qquad (2)$$

It should be noted that the reason for putting $1 - p_i$ rather than p_i in the numerator of equation (1) is that it is desirable to make the variables z_i monotonically increasing in the magnitude of the factors f_{ij} rather than monotonically decreasing. For example, the magnitude of the largest number in a problem increases with the difficulty of the problem, and it is desirable that the model reflect this increase in a direct rather than in an inverse fashion.

In the case of latencies a transformation like (1) is not required. Let t_i be the mean latency on problem-item i for a given group of subjects. We then apply the same model as (2), namely,

$$t_i = \sum_j \beta_j f_{ij} + \beta_o . \qquad (3)$$

It is also evident that no transformation is required to make latencies monotonically increasing in the expected difficulty of the factors. We have shown different weights β_j for the latencies, because the empirical interpretation of the weights must necessarily be different for the variables z_i, but as we would expect, there is a high positive correlation between the weights α_j and β_j. It is worth noting that in the case of the analysis of the latencies, the individual factors and their weights may be identified as the direct contribution of a given factor to the total latency. Thus, for example, the contribution of factor j to the total latency is just the number $\beta_j f_{ij}$ which is scaled in seconds. The constant that arises in the linear-regression model may be interpreted as the constant orientation and preparation time required in solving the problems of the class under investigation.

The variables we consider are of two sorts. The first is the kind of 0,1 - variable standard in the analysis of variance. Such a variable would be appropriate, for example, in dealing with problem format. The second kind of variable is one that is in principle continuous, although in practice it assumes a finite set of values for the problems being considered here. For most of these variables the conception and formal definitions of the variables are quite straightforward within the context of elementary arithmetic itself. Typical variables have already been mentioned; however, the variable or factor dealing with the number of steps required to solve a problem is most important from the standpoint of the psychological analysis. This factor also seems most promising for future developments of the models presented in this paper. We turn now to the appropriate formal definitions. As has already been emphasized, we feel that the greatest possibilities for subsequent theoretical analysis lie in this direction. What we report here is only the result of our first relatively crude analysis, and we are already heavily engaged in the process of deepening this analysis, particularly by breaking up the single variable of number of steps into several components. Some preliminary results are reported at the end of the paper.

The steps postulated have been broken up into three classes: those required to transform the problem into canonical form, those corresponding to the number of operations performed, and those corresponding to the number of digits that must be held in memory. We refer to these three classes as the transformation, operation, and memory classes. As will be seen, there is a quite high correlation in most problems between the number of operation steps and the number of memory steps. An essential point for later work is to make these two processes more orthogonal in operational characterization. Another assumption that is surely too simple is reflected in the assignment of the same weight to addition and subtraction, in the analysis of operation steps. Other unrealistic simplifications have been made, but the general definitions required to characterize the number of steps required for solution are still relatively complex, and we think they constitute a reasonable beginning.

For simplicity we first consider just the transformation steps that convert any problem into canonical form. By canonical form we mean the equational form in which the blank or unknown stands by itself as the only term to the right of the equal sign. Thus for numbers m, n and p, regardless of whether the numbers are one digit or two digit, we have

(i) $m \pm n = \_\_$ is already in canonical form,

(ii) $m + \_\_ = p$ is transformed to $\_\_ = p - m$,

which is transformed to $p - m = \_\_$, requiring two steps,

(iii) $\_\_ + n = p$ is identical to (ii)

(iv) $m - \_\_ = p$ is transformed to $m - p = \_\_$, requiring one step,

and finally

(v) $\_\_ - n = p$ is transformed to $n + p = \_\_$, also requiring one step. The fact that $m + \_\_ = p$ requires one more transformation than $m - \_\_ = p$ agrees with the intuition that (ii) is really more difficult than (iv). We make explicit the number T of transformations in the following definitions that formalize (i) - (v).

$$T(m \pm n = \_\_) = 0$$

$$T(m + \_\_ = p) = 2$$

$$T(\_\_ + n = p) = 2$$

$$T(m - \_\_ = p) = 1$$

$$T(\_\_ - n = p) = 1$$

$$T(m + n = p + \_\_) = 1$$

$$T(m + n = \_\_ + p) = 1$$

The last two equations cover two additional cases that arise in the data we analyze.

Turning now to the operation and memory steps, we need to make explicit the number of digits involved, so we always use initial letters of the alphabet for single digits. Also, because we postulate that the operation and memory steps enter only after the transformation to canonical form has taken place, we may simplify the notation, writing, for example, $O(ab + cd)$ or $M(ab + cd)$ for the number of operation or memory steps respectively. For example,

$$O(5 + 0) = 0$$

but

$$O(5 + 4) = 1$$

because we postulate no operation is required for handling zero.

$$O(15 + 12) = 2$$

because one operation is $5 + 2$ and the second is $1 + 1$. On the other hand, in the form $ab + cd$, when $b + d > 9$, there are three operations. For example,

$$O(25 + 47) = 3 ,$$

because one operation is $5 + 7$; the second is the partial sum $1 + 2$ using the 1 that is "carried"; and the third is $3 + 4$, the partial sum plus 4, the other tens' digit.

In the case of memory,

$$M(15 + 12) = 1 ,$$

because only 7, the sum of 5 and 2, must be held in memory while the tens are added and the correct tens' digit response is made (the problem format required input of the tens' digit before the ones' digit). On the other hand,

$$M(25 + 47) = 3$$

because (i) the 2 of 12, the sum of 7 and 5, must be held in memory for the ones' response, (ii) the 1 which is carried to the tens' place must be held, and (iii) the partial sum $1 + 2$ must be held while it is added to 4. The definition for the more complicated format $ab + cd - ef$ is given recursively in terms of $ab + cd$, and thus does not need a separate treatment. Formally the definitions of the number of operation and memory steps are as follows:

$$O(a + b) = \begin{cases} 0 & \text{if } a = 0 \text{ or } b = 0 \\ 1 & \text{if } a \neq 0 \ \& \ b \neq 0 \end{cases}$$

$$O(ab + d) = \begin{cases} O(b + d) & \text{if } b + d \leq 9 \\ O(b + d) + 1 & \text{if } b + d > 9 \end{cases}$$

$$O(ab + cd) = \begin{cases} O(b + d) + 1 & \text{if } b + d \leq 9 \\ O(b + d) + 2 & \text{if } b + d > 9 \end{cases}$$

$$O(a - b) = \begin{cases} 0 & \text{if } b = 0 \\ 1 & \text{if } b \neq 0 \end{cases}$$

$$O(ab - c) = \begin{cases} O(b - c) & \text{if } b \geq c \\ O(b - c) + 1 & \text{if } b < c \end{cases}$$

$$O(ab - cd) = \begin{cases} 1 & \text{if } d = 0 \\ 2 & \text{if } b \geq d > 0 \\ 3 & \text{if } b < d \end{cases}$$

$$M(a \pm b) = 0$$

$$M(ab + c) = \begin{cases} 1 & \text{if } b + c \leq 9 \\ 2 & \text{if } b + c > 9 \end{cases}$$

$$M(ab + cd) = \begin{cases} 1 & \text{if } b + d \leq 9 \\ 3 & \text{if } b + d > 9 \end{cases}$$

$$M(ab - c) = \begin{cases} 1 & \text{if } b \geq c \\ 2 & \text{if } b < c \end{cases}$$

$$M(ab - cd) = \begin{cases} 1 & \text{if } b \geq d \\ 3 & \text{if } b < d\text{ .} \end{cases}$$

If $ab + cd = gh$ then

$$O(ab + cd - ef) = O(ab + cd) + O(gh - ef)$$

and

$$M(ab + cd - ef) = M(ab + cd) + M(gh - ef) + 1 .$$

The additional step in the case of $M(ab + cd - ef)$ comes in from having to remember $a + c$, or $a + c + 1$, as the case may be, which is not part of $M(ab + cd)$ or $M(gh - ef)$.

Similarly, if $ab + c = fg$ then

$$O(ab + c - de) = O(ab + c) + O(fg - de)$$

and

$$M(ab + c - de) = M(ab + c) + M(fg - de) .$$

A corresponding definition holds for $ab + cd - e$.

In evaluating problem structure, we determine the total number of transformation, operation, and memory steps. Thus, for example,

$$25 + 26 = 18 + \_\_$$

has the maximum number of 14 steps, because

$$T(25 + 26 = 18 + \_\_) = 1$$
$$O(25 + 26 - 18) = 6$$
$$M(25 + 26 - 18) = 7 ,$$

and on the other hand the problem

$$5 + 0 = \_\_$$

has the minimum of 0 steps. Of course, some students will solve many individual problems by a shorter method, and the present approach to counting steps does not incorporate any such special methods. This again is a matter for subsequent investigation.

In the analyses reported in this paper we have entered the total number N of steps as a single variable for most of the results reported, but in one case we have broken the steps up into classes, and further intensive work in this direction is currently underway. In the linear-regression models used for this purpose, we replace αN by $\alpha_1 T + \alpha_2 O + \alpha_3 M$.

3. Method.

The data reported and analyzed in this paper were collected as an integral part of a full academic-year, operational program in computer-assisted mathematics instruction. For this reason we shall describe in some detail this program.

Subjects. The approximately 270 subjects in this project consisted of the entire population of grades three, four, five, and six in the Grant Elementary School, except for those in the handicapped classes.

The children came from a middle-class, suburban community. All children lived within walking distance of the school.

Although there was some fluctuation in attendance figures during the year, school records show the following population figures at year's end. There were 32 boys and 30 girls in grade three, 41 boys and 35 girls in grade four, and 44 boys and 26 girls in grade five. The mean I.Q. of the fifth-grade group was 114, the range 72-145. Grade six had 35 boys and 27 girls. Mean I.Q. of the sixth-grade class was 117, range 88-156. There were no data on I.Q. scores for either grade three or grade four.

Equipment. The student terminals used in this project were commercially available teletype machines, connected by private, high-speed, phone lines to the Institute's computer at Stanford. A large book closet, which opened into the classroom, was modified by adding a ventilation fan, light, and electrical outlet. This provided privacy for the user and insulated the rest of the class from the operational noise of the teletype.

The control functions for the entire system were handled by a medium-sized computer. The PDP-1 has a 16,000 word core, and a 4,000 word core which can be interchanged with any of 32 bands of a magnetic drum. All input-output devices are processed through a time-sharing system. Two high-speed data channels permit simultaneous computation and servicing of peripheral devices. Additional backup in computational power, additional storage, and increased input-output speed are obtained through connections to disk storage of a larger computer (IBM 7090) located at the Stanford Computation Center.

Response time was measured from the instant (nearest .001 sec.) the type wheel was in position at the response area (or answer blank). When the student depressed one of the keys on the teletype's keyboard, a signal was sent to the computer. The character was recognized by the computer approximately one millisecond (.001 sec.) after being initiated by the student. A reading was taken from a real-time clock, internal to the computer, and this information compared to the time read when the type wheel was positioned. Under optimal conditions latency measurements could be made with an accuracy of from two to three milliseconds. However, as mentioned above, the system was operating under a time-sharing arrangement. This reduced the level of accuracy of latency measure to about one-tenth of a second. Conversion from readings in thousandths to the nearest tenth was made by division and truncation.

__Curriculum Materials__. Daily lessons were prepared and organized by concepts or topics into blocks or units. The concept blocks were arranged sequentially corresponding approximately to the order of topics in the textbook series, __Sets and Numbers__, written by the first author of this paper. The length of time needed to complete a concept block varied from three to twelve days, when a single lesson was taken each day. The curriculum objective of the daily lessons was to provide an organized program of review, maintenance and drill on basic skills and concepts of elementary mathematics, particularly arithmetic. Initial instruction in all concepts was given initially by the teacher, and consequently the drill-and-practice work at computer terminals did not include a detailed introduction to the concepts.

Teachers involved in the project were free, subject to certain constraints, to select any of the prepared blocks in order to correlate the drill-and-practice work with their daily instruction. Handbooks were furnished which described available concept blocks in detail. Also included in the handbooks were reprints of every lesson. Table 1 describes the concept blocks prepared and planned for each grade level.

Insert Table 1 about here

Each concept block was organized in the manner shown diagrammatically by Figure 1. Lessons were prepared at each of five levels of difficulty within each concept block. Among factors which determined intuitive estimates of relative difficulty are those discussed in this paper. They, and the exercises reflecting them, were chosen intuitively on the basis of teaching experience and previous project experience gained from preparation and testing of the textbook series cited above. Each class was restricted to a single concept block at a time. On the first day of a new block, every member of a class was given the same lesson. This lesson was of average difficulty (level 3). Those students who scored between 60% and 79% were given a level-three lesson the following day;

Insert Figure 1 about here

----

those who scored above 79% were given a lesson on the next higher level (level 4); those who failed to score at least 60% were given a simpler lesson on a lower level (level 2). This procedure was followed throughout a concept block, that is, a score of above 79% branched a student up one level each day, while a score of below 60% branched a student down one level each day, but of course a student could not move up beyond level 5 or down below level 1. Thus, by day three, a student could have been at any one of five levels, with a different lesson at each level. It was intended that approximately 90% of the students would alternate between levels two, three, and four, and that those remaining on any level would be nearly homogeneous. Level 1 was mainly remedial in character, and level 5 was ordinarily meant to be difficult. Drills on all levels increased somewhat in difficulty from day to day within a block as successively more advanced aspects of each topic were reviewed.

Program Logic. Under computer control each problem was completely typed out, including a blank for the response. The type wheel of the teletype was then positioned at the blank so that the response would be properly placed. A correct response was reinforced by the appearance of the next exercise. When an incorrect first response was made, the word "wrong" was typed out and the exercise itself was retyped. A second error on the same exercise was followed by the message "wrong, the answer is _ _", with the correct answer being displayed. The exercise itself was then retyped once more to allow for a correction response. An error

Table 1

Concept blocks for grades 3-6. Number of days spent on each block is shown.

	Grade 3		Grade 4		Grade 5		Grade 6	
Blocks	**Days**	**Description**	**Days**	**Description**	**Days**	**Description**	**Days**	**Description**
1	8	sums 0-20	5	sums 0-40	5	sums 11-60	5	mixed drill, all operations
2	5	differences 0-20	5	differences 0-40	5	differences 11-60	10	fractions, addition, subtraction, changing terms
3	3	mixed addition and subtraction	5	sums 31-70	10	multiplication tables 3-12	5	multiplication tables 2-12
4	5	multiplication tables 2 and 3	5	differences 31-70	5	multiplication tables, c=axb	4	factors, multiples and primes
5	4	mixed addition and subtraction	7	multiplication tables 4-10	5	long division	5	fractions, simple equations
6	5	word problems* mixed review	5	division, 6-12 tables, a/b=c	7	mixed drill, all operations, word problems	10	achievement tests
7	3	mixed review, addition, subtraction, multiplication	5	mixed review, all operations, inequalities	10	fractions	8	factors, multiples, fractions
8	10	addition with carrying	8	CAD laws for multiplication, addition, subtraction	5	units of measure	5	multiplication of large numbers
9	10	subtraction, regrouping	5	multiplication table 6-12	6	CAD laws	5	word problems, all operations
10	3	money, equivalence*	5	mixed drill, all operations	5	using CAD laws, giving reasons	3	long division, standard form, 1-digit divisor
11	5	mixed review	5	word problems	3	word problems	8	decimal and common fractions, per cent*

Table 1 (continued)

	Grade 3		Grade 4		Grade 5		Grade 6	
Blocks	Days	Description	Days	Description	Days	Description	Days	Description
12			4	CAD** Laws* addition, subtraction, multiplication	5	CAD Laws, giving reasons	4	metric units of measure*
13			6	distribution law for division*, CAD Laws	9	mixed drills, all operations	10	mixed drills
14	5	multiplication tables 0-5	5	mixed drills, all operations	5	fractions* word problems	4	logic
15			5	subtraction, fractions	5	division	5	fractions*, addition, subtraction, multiplication, division
16			5	fractions*, addition, subtraction	2	addition and subtraction of decimals	5	decimal operations*
17			5	multiplication	5	addition, subtraction of integers	4	word problems*
18			10	multiplication	6	exponents, word problems*	8	mixed review, all operations*
19			15	mixed multiplication, subtraction*	2	metric measure	5	CAD Laws
20			15	mixed multiples of 10*	10	mixed drill, all operations	10	long division, ladder form, 1- and 2-digit divisors
21			2	fractions, addition, subtraction*	8	multiplication-exponents*	4	metric units of measure
22			5	mixed drills*, all operations	3	coordinate systems	5	mixed review, negatives, inequalities*
23			5	word problems	9	sets, review*	8	word problems giving reasons*
24	5	multiplication tables	5	negatives, addition, subtraction*	5	word problems*	10	long division, ladder form, 1- and 2-digit divisors

Table 1 (continued)

	Grade 3		Grade 4		Grade 5		Grade 6	
Blocks	Days	Description	Days	Description	Days	Description	Days	Description
25	7	3- and 4-digit column addiiton	5	mixed addition, subtraction*	5	CAD Laws giving reasons	5	mixed review*
26			5	CAD Laws*	10	logic	5	vertical subtraction
27			3	mixed multiplication, division*	5	per cent	2	special addition
28			5	word problems, units of measure*	10	achievement tests	2	special multiplication
29			4	mixed review, all operations*	5	vertical subtraction	12	long division, ladder form, 1-digit divisor
30			2	CAD Laws*	2	special addition	7	long division, standard form, 2-digit divisor
31	10	multiplication to 12 × 12 vertical form	4	word problems, units of measure*	2	special multiplication	7	long division, ladder form, 2-digit divisor
32	5	CAD Laws for addition, subtraction, multiplication	10	achievement tests	5	long division, ladder form, 2-digit divisor	2	division tests, multiple-choice form, basic concepts
33	5	division facts to 12 × 12*	5	column subtraction				
34	5	3-4 digit column addition with regrouping*	2	special addition				
35	5	mixed review	2	special multiplication lessons				
36	2	special addition	10	remedial multiplication tables 3-7				

Table 1 (continued)

	Grade 3		Grade 4		Grade 5		Grade 6	
Blocks	Days	Description	Days	Description	Days	Description	Days	Description
37	2	special multiplication	10	remedial multiplication tables 4-9				
38		column multiplication*	5	division with variables standard form				
39			7	long division, ladder form, 1-digit divisor				
40			7	long division, ladder form, 1-digit divisor				
41			5	fractions				

*Blocks planned but not written.

**CAD stands for commutative, associative and distributive.

on the correction response caused the correct answer to be given again, but whether the third response was correct or incorrect, the next exercise was presented.

If a response was not given within a predetermined interval of time, usually ten seconds, the machine response followed the above pattern except that the words "time is up" were substituted for the word "wrong" at each step described above. A flow chart of the program logic is given in Figure 2.

Insert Figure 2 about here

Procedure. The two classes in each of grades four, five, and six began in October, 1965 sharing one teletype between them. One class was scheduled to run in the morning, the other in the afternoon. However, this proved to be an unworkable arrangement. Beginning with the third week, the classes worked on the teletype on alternate days. The machines were operated daily between the hours of 8:30 A.M. and 3:00 P.M.

In late February, 1966, the two third-grade classes began daily lessons with the addition of two more teletypes, which brought the total number of machines in operation on a daily basis to five. In early April, 1966, the last three machines were put in operation, bringing the total number of teletypes to eight. Each class in grades three, five, and six had its own teletype. Grade four had been divided into three classes to alleviate an overcrowded situation. One of the fourth-grade classes had its own machine, the other two classes shared the remaining teletype.

The students took their lessons one at a time on each machine in the order prescribed by their teacher. The program began by asking the student

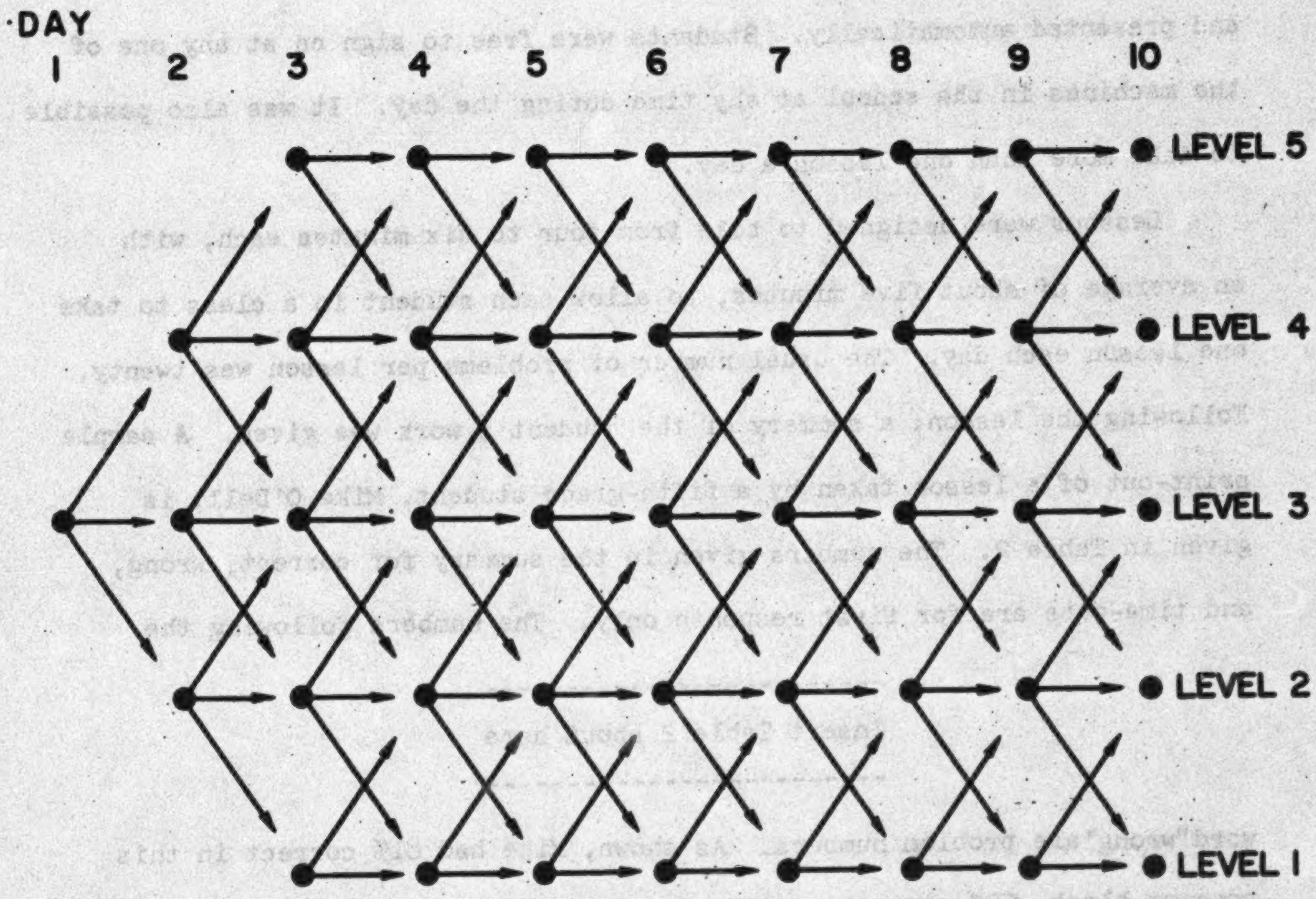
DAY
1
2
3
4
5
6
7
8
9
10
LEVEL 5
LEVEL 4
LEVEL 3
LEVEL 2
LEVEL 1

Figure 1. Diagram of branching structure followed in constructing sets of exercises for concept blocks.

to type his name. When the name had been correctly input, the lesson began as described above in the section on program logic. If a student failed to spell his name correctly, or gave a fictitious name (such as Batman), the program asked him to try again. An individual history was kept in computer memory for each student. When a student's name was input correctly, the proper lesson was selected, based on the branching criteria, and presented automatically. Students were free to sign on at any one of the machines in the school at any time during the day. It was also possible to take more than one lesson a day.

Lessons were designed to take from four to six minutes each, with an average of about five minutes, to allow each student in a class to take one lesson each day. The usual number of problems per lesson was twenty. Following the lesson, a summary of the student's work was given. A sample print-out of a lesson taken by a fifth-grade student, Mike O'Dell, is given in Table 2. The numbers given in the summary for correct, wrong, and time-outs are for first response only. The numbers following the word "wrong" are problem numbers. As shown, Mike had 81% correct in this concept block, 59% correct to date for the whole school year, which began October 18 on the teletype. The time given in hours, minutes, and seconds is the total time Mike had spent on all teletype lessons to date.

Insert Table 2 about here

The students were not allowed to use pencil or paper when working on the teletype. Each exercise was worked on the machine so that all responses could be recorded and latencies measured. The response mode

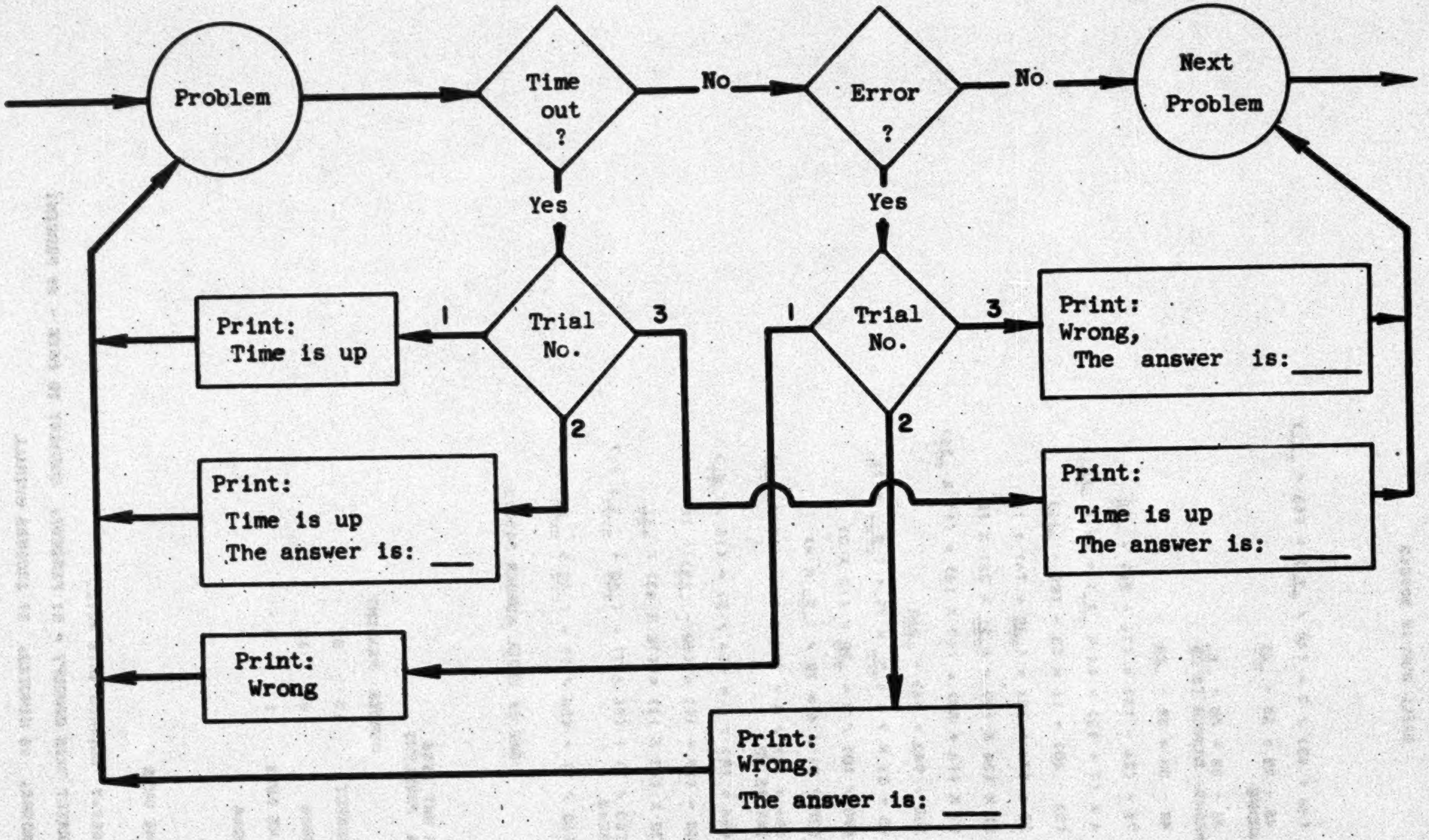
Problem
Time out ?
No
Error ?
No
Next Problem
Yes
Yes
Trial No.
1
2
3
Trial No.
1
2
3
Print: Time is up
Print: Time is up The answer is: ___
Print: Wrong
Print: Wrong, The answer is: ___
Print: Wrong, The answer is: ___
Print: Time is up The answer is: ___

Figure 2. Flow chart of the program logic for presentation of problems and classification of responses.

Table 2. Sample print out of a fifth-grade student.

```
PLEASE TYPE YOUR NAME


MIKE ODELL

          DRILL NUMBER 509013


 (42 + 63) / 7 = (42 / _7_) + (63 / _7_)

 48 - 38 = 38 - _48_
WRONG

 48 - 38 = 38 - _4_
WRONG, ANSWER IS 28

 48 - 38 = 38 - _28_

 76 - (26 - 10) = (76 - 26) + _10_

 4 X (7 + 13) = (4 X _7_) + (4 X _13_)

 (53 - 20) - 11 = 53 - (20 + _11_)

 38 + (74 + 18) = (_38_ + 74) + 18

 51 X (36 X 18) = (_51_ X 36) X 18

 17 X (14 + 34) = (17 X 14) + (17 X _34_)

 368 + 943 = 943 + _368_

 (5 + 8) X 7 = (_5_ X 7) + (_8_ X 7)

 (90 / 10) / 3 = _90_ / (10 X 3)

 (72 / 9) / 4 = 72 / (_9_ X 4)

 (54 + 18) / 6 = (54 / 6) + (18 / ___)
TIME IS UP

 (54 + 18) / 6 = (54 / 6) + (18 / _6_)

 60 - (19 - 18) = (60 - _19_) + 18

 72 X (43 X 11) = (72 X 43) X _11_

 (63 / 7) + (56 / 7) = (_63_ + _7_) / 7
WRONG

 (63 / 7) + (56 / 7) = (_63_ + _56_) / 7

          END OF DRILL NUMBER 509013


 13 MAY 1966
 16  PROBLEMS

          NUMBER  PERCENT

 CORRECT      13      81

 WRONG         2      12

 TIME OUTS     1       6

 WRONG
 2
 16


 TIME OUTS
 13
   222.7  SECONDS THIS DRILL

 CORRECT THIS CONCEPT - 81 PERCENT.  CORRECT TO DATE - 59 PERCENT

 4 HOURS.  46 MINUTES.  59 SECONDS OVERALL


 GOODBYE MIKE.
```

was limited to either numerical answers or simple single-letter answers for multiple-choice problems.

Initial instruction on the teletype and program operation consisted of explaining to each class the general procedure of taking turns on the machine, and of showing that only the answer need be input on the keyboard. The program logic was also explained. Staff members helped each student find the letters to type his name for the first two or three lessons. Students had little trouble learning how to type their names or answer the questions.

Following the summary and "goodbye" message the student was told "please tear off on dotted line". A dotted line was printed, and the student then tore off his print-out and took it with him as a permanent record of his work.

4. Results.

To begin with, it must be emphasized that we have not attempted a detailed model-theoretic analysis of data from all the concept blocks listed in Table 1. We have selected five topics on which we had considerable data and which were sufficiently simple to provide a good starting point. The first analysis deals with fourth-grade and fifth-grade performance on addition; the data are drawn from blocks 1 and 3 of grade four and block 1 of grade five listed in Table 1. The second analysis is concerned with subtraction at the same grade levels; the data are drawn from block 2 for each grade. The third analysis looks at fourth-grade multiplication data, drawn from block 5. The fourth analysis deals with a relatively-controlled experiment on the multiplication tables for grades 3-6; the data are drawn from blocks 37, 35, 31, and 28, respectively, for each of grades 3-6. The final analysis returns to the data of the first analysis and looks at the results of breaking up the regression analysis of the number of steps into several variables, as indicated in the theoretical discussion.

As remarked earlier, for each set of problems examined success-latency and error probability have been treated as separate dependent variables. Separate regression coefficients were obtained from the same independent variables to predict latency and error probability. This is justifiable by the intuitive assumption that success-latency and error probability are different measures of common underlying processes, and is justified empirically by our finding that the correlation between the two dependent variables was consistently greater than 0.7 for the data we have collected.

To minimize the effects of subject variables such as I.Q., the problems and data were usually treated separately by grade, concept block and level, as is made explicit in the tables given below. It is assumed that children working within a given branching level form a more homogeneous group of subjects than children working on different level problems. We were unable to analyze data from some of the levels available because too few children entered those branches.

The first step in analysis was to obtain regression coefficients for each grade and level for the two dependent variables. A stepwise, multiple linear-regression analysis program, BIMD 02R, adapted for Stanford University's IBM 7090 computer, was used to obtain regression coefficients, multiple correlation R and R^2. For a finer-grained analysis of the goodness of fit of the success-latency predicted from the regression model and observed success-latency, a computer program was written to calculate the predicted mean success-latency for each problem and to give as a measure of fit

$$S^2 = \frac{\sum_{i=1}^{N} (\text{obtained latency}_i - \text{predicted latency}_i)^2}{N - k}$$

where N is the number of problems for which the latency was predicted and k is the number of estimated parameters. Similarly, for a finer analysis of the goodness of fit of the regression model to the error data, a program was written to calculate the predicted proportion of errors for each problem i from the obtained regression coefficients and to give as a measure of fit X_i^2, where

$$X_i^2 = \frac{(f_i - p_i N)^2}{p_i(1 - p_i)N}$$

and

f_i = observed frequency of correct responses,
p_i = predicted probability of a correct response,
N = number of students.

__Addition--grades four and five__. The three independent variables used in the regression analyses for addition were the variable NSTEPS, which was described in detail earlier, and the two magnitude variables, magnitude of sum (MAGSUM) and magnitude of the smallest addend (MAGSMALL). It is obvious that the value of MAGSUM and MAGSMALL is independent of whether the problem for the student was to find the missing sum or a missing addend. For example, in the three related problems $7 + 9 = \_\_$, $7 + \_\_ = 16$ and $\_\_ + 9 = 16$, MAGSUM = 16 and MAGSMALL = 7.

The coefficients obtained for the regression equations are shown in Table 3. This table indicates the level of problems analyzed, (Level), the number of children who worked on the problems

----

Insert Table 3 about here

----

in that level, (Subjects)[2], the number of different problems analyzed, (Problems)[3], the regression constant and the regression coefficients for the three independent variables. The absence of a value of a given coefficient indicates that the variable it applies to made no significant contribution to the regression equation, and the computer program therefore did not use that variable in obtaining a regression line. In reading the regression table it should be remembered that the transformation described previously was applied to the observed proportion of errors, and therefore when obtaining a prediction from the coefficients for proportion of errors, the numbers z_i calculated from the coefficients must be transformed to obtain the predicted proportion of errors.

It is clear from scanning the coefficients in Table 3 that NSTEPS is the most important of the three variables in predicting both errors and success-latencies. A rough indication of the goodness of fit of the regression lines is reflected by the multiple-correlation coefficient R

Table 3

Linear-regression coefficients for fourth- and fifth-grade addition

Grade 4 Addition, Set 1, Proportion of Errors

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
2	6	19	-2.73	0.16	0.09	-0.03	0.61	0.37
3	21	38	-2.65	0.16	0.05	-0.03	0.56	0.32
4	24	38	-1.44	0.24	-0.01	0.05	0.86	0.74
5	9	19	-1.74	0.08	0.03	-0.09	0.40	0.16

Grade 4 Addition, Set 1, Success Latency

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
2	6	19	0.24	0.18	0.14	-0.07	0.64	0.40
3	21	38	-0.76	0.47	0.13	-0.09	0.69	0.48
4	24	38	2.32	0.57	-0.02	0.07	0.86	0.74
5	9	19	2.19	0.17	0.00	0.00	0.44	0.19

Grade 4 Addition, Set 2, Proportion of Errors

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
2	7	57	-1.69	0.17	0.02	-0.02	0.54	0.29
3	41	95	-0.73	0.21	-0.01	-0.01	0.64	0.41
4	34	76	-1.60	0.20	0.00	0.01	0.80	0.64

Grade 4 Addition, Set 2, Success Latency

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
2	7	57	0.95	0.56	0.06	-0.09	0.64	0.42
3	41	95	1.77	0.73	0.01	-0.06	0.82	0.68
4	34	76	1.55	0.47	0.01	0.02	0.75	0.56

Grade 5 Addition, Proportion of Errors

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
3 and 4 combined	12	57	-2.41	0.10	0.03	0.03	0.81	0.66

Grade 5 Addition, Success Latency

Level	Subjects	Problems	Constant	NSTEPS	MAGSUM	MAGSMALL	R	R^2
3 and 4 combined	12	57	-2.22	0.47	0.09	0.07	0.73	0.54

and its square (R^2) which is an estimate of the amount of variance accounted for by the regression model. In only one case is less than 40% of the success-latency variance accounted for by the model. When one takes into account that the two magnitude variables account for a relatively small amount of the variance, and that in setting up the variable NSTEPS we have combined several potentially powerful and probably independent variables, the results are encouraging.

Tables 4-10 present the X^2 and individual contributions of the problems to X^2 when the seven sets of coefficients for response errors given in Table 3 were used to predict the proportions of errors. Included in these tables are the rank order of observed problem difficulty, the observed proportion of students making errors, (Observed $(1 - p_i)$); the proportion of errors predicted from the linear-regression model, (Predicted $(1 - p_i)$); and the actual component of the X^2 contributed by the problem.

Insert Tables 4, 5, 6, 7, 8, 9, 10 about here

Table 11 presents the same analysis for fifth-grade addition. Unfortunately the data on addition for the fifth-grade children are rather sparse.

Insert Table 11 about here

The fifth graders were presented with a concept block on addition during the first week of operation of the computer-based system. Technical difficulties and initial student unfamiliarity with the teletypes caused us to lose a good part of the data we had hoped to collect.

Table 4

Predicted and observed proportions of errors and success-latency in fourth-grade addition, concept block 1, level 2

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
1	24 + __ = 24	0.08	0.21	2.40	2.65	0.61
2	24 + 3 = __	0.17	0.44	4.50	4.05	1.77
3	__ + 0 = 21	0.17	0.22	4.70	3.46	0.10
4	26 + __ = 26	0.17	0.26	3.00	2.78	0.28
5	__ + 0 = 23	0.17	0.30	3.00	3.73	0.49
6	26 + 2 = __	0.33	0.50	3.70	4.26	0.71
7	26 + __ = 27	0.33	0.30	2.50	2.92	0.02
8	23 + __ = 26	0.33	0.31	2.40	3.00	0.02
9	22 + __ = 30	0.50	0.69	5.50	3.98	0.98
10	21 + 7 = __	0.50	0.42	4.80	3.90	0.17
11	24 + __ = 29	0.50	0.43	3.80	3.33	0.12
12	29 + __ = 30	0.50	0.57	3.30	3.47	0.12
13	22 + 8 = __	0.50	0.67	4.80	4.46	0.80
14	__ + 2 = 25	0.67	0.53	4.90	4.21	0.45
15	21 + 8 = __	0.67	0.45	3.70	3.96	1.16
16	__ + 8 = 30	0.67	0.81	4.30	4.82	0.76
17	23 + __ = 28	0.83	0.40	2.90	3.27	4.72
18	__ + 3 = 24	0.83	0.46	3.30	4.01	3.32
19	__ + 7 = 28	0.83	0.59	3.00	4.26	1.43

$\chi^2 = 18.03$ (19 items) χ^2(items < 10) = 18.03 (19 items) $s^2 = 0.65$

Table 5

Predicted and observed proportions of errors and success-latency in fourth-grade addition, concept block 1, level 3

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
1	26 + __ = 29	0.02	0.28	3.40	5.16	7.96
2	22 + __ = 24	0.02	0.18	2.90	4.58	4.27
3	24 + 0 = __	0.02	0.04	2.10	2.41	0.24
4	26 + 0 = __	0.02	0.05	2.30	2.67	0.45
5	36 + 0 = __	0.03	0.16	2.30	3.99	2.22
6	31 + 0 = __	0.03	0.09	2.00	3.33	0.86
7	37 + __ = 39	0.03	0.10	2.80	3.58	1.00
8	40 + __ = 40	0.03	0.05	2.50	2.51	0.13
9	26 + 1 = __	0.04	0.11	3.40	3.66	1.02
10	37 + __ = 40	0.06	0.20	3.60	4.65	2.41
11	24 + __ = 27	0.08	0.23	4.40	4.89	3.02
12	22 + 1 = __	0.08	0.07	2.70	3.13	0.09
13	22 + __ = 25	0.08	0.19	3.30	4.63	1.84
14	35 + __ = 40	0.11	0.23	3.00	4.82	1.40
15	33 + __ = 39	0.11	0.13	4.00	3.92	0.04
16	27 + 0 = __	0.13	0.06	3.10	2.81	2.01
17	23 + __ = 27	0.13	0.22	3.70	4.81	1.27
18	21 + __ = 30	0.17	0.37	4.80	5.72	4.21
19	31 + __ = 34	0.17	0.08	4.40	3.43	1.61
20	__ + 1 = 37	0.17	0.45	4.50	5.92	5.94
21	33 + 6 = __	0.28	0.15	4.30	4.02	2.56
22	__ + 2 = 23	0.29	0.12	5.70	3.98	6.45
23	__ + 2 = 26	0.29	0.17	5.30	4.38	2.68
24	23 + 7 = __	0.33	0.18	6.60	4.48	3.49
25	__ + 3 = 35	0.33	0.36	5.50	5.48	0.05
26	32 + 4 = __	0.33	0.15	4.30	4.06	4.86
27	31 + __ = 38	0.33	0.13	4.90	3.96	6.54
28	__ + 4 = 38	0.39	0.43	5.70	5.79	0.12

Table 5 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
29	32 + __ = 39	0.44	0.14	5.20	4.00	14.57
30	__ + 1 = 32	0.44	0.31	6.30	5.26	1.51
31	26 + 3 = __	0.46	0.11	4.50	3.75	27.76
32	__ + 5 = 30	0.54	0.35	5.70	5.59	3.77
33	__ + 6 = 39	0.56	0.42	6.10	5.75	1.31
34	__ + 6 = 28	0.58	0.16	5.90	4.30	31.77
35	__ + 9 = 30	0.63	0.29	5.40	5.24	13.19
36	__ + 6 = 27	0.67	0.14	5.50	4.17	52.85
37	__ + 5 = 37	0.67	0.38	7.20	5.58	6.20
38	__ + 4 = 40	0.83	0.67	8.10	6.99	2.27

$X^2 = 233.91$ (38 items) X^2(items < 10) $= 83.78$ (33 items) $S^2 = 1.29$

Table 6

Predicted and observed proportions of errors and success-latency in fourth-grade addition, conceptblock 1, level 4

Rank	Equations	Observed $(1 - p_1)$	Predicted $(1 - p_1)$	Observed Latency	Predicted Latency	X^2
1	0 + 29 = 0 + __	0.01	0.03	2.40	2.23	0.40
2	1 + 38 = __ + 0	0.05	0.13	3.20	3.72	0.69
3	1 + 27 = 0 + __	0.06	0.16	2.60	3.97	2.84
4	2 + 36 = 0 + __	0.09	0.13	3.80	3.74	0.16
5	0 + 34 = 0 + __	0.09	0.03	3.40	2.12	1.45
6	3 + 26 = 0 + __	0.14	0.16	3.70	3.95	0.08
7	1 + 25 = __ + 0	0.17	0.16	2.90	4.02	0.00
8	4 + 27 = 2 + __	0.27	0.68	7.30	6.33	8.30
9	33 + 3 = __ + 29	0.27	0.79	5.70	6.86	17.51
10	12 + 10 = __ + 0	0.28	0.18	4.50	4.11	2.53
11	10 + 16 = 6 + __	0.39	0.41	3.90	5.00	0.04
12	27 + 2 = 10 + __	0.39	0.64	5.20	5.78	9.75
13	7 + 18 = 0 + __	0.39	0.38	4.50	5.19	0.01
14	9 + 28 = 7 + __	0.46	0.77	6.50	6.53	5.91
15	10 + 29 = 8 + __	0.46	0.53	5.00	5.41	0.26
16	11 + 12 = __ + 1	0.53	0.56	5.80	5.87	0.16
17	24 + 3 = 5 + __	0.58	0.46	5.00	5.35	2.37
18	17 + 5 = __ + 11	0.58	0.40	6.20	5.02	5.04
19	9 + 14 = __ + 2	0.58	0.72	7.50	6.51	3.11
20	9 + 18 = __ + 5	0.61	0.76	7.40	6.63	4.68
21	34 + 5 = 11 + __	0.64	0.59	5.90	5.78	0.11
22	22 + 7 = __ + 14	0.72	0.69	7.60	6.15	0.19
23	7 + 22 = 6 + __	0.72	0.53	6.00	5.50	5.37
24	11 + 28 = 8 + __	0.73	0.67	6.00	5.98	0.19
25	27 + 7 = __ + 20	0.73	0.66	6.40	6.03	0.19
26	17 + 22 = 5 + __	0.73	0.59	7.00	5.78	0.90
27	35 + 3 = __ + 12	0.73	0.54	---	---	1.62
28	30 + 2 = __ + 5	0.73	0.54	4.60	5.74	1.54

Table 6 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
29	23 + 2 = __ + 8	0.75	0.83	7.00	6.88	1.46
30	25 + 4 = 11 + __	0.75	0.61	5.70	5.94	2.90
31	19 + 8 = __ + 6	0.75	0.78	6.40	6.70	0.25
32	32 + 5 = __ + 9	0.91	0.72	---	---	1.91
33	29 + 7 = __ + 15	0.91	0.85	6.10	7.13	0.26
34	22 + 12 = 16 + __	0.91	0.95	8.10	8.09	0.39
35	12 + 22 = __ + 6	0.91	0.85	---	---	0.34
36	33 + 1 = 7 + __	0.91	0.64	7.60	6.19	3.49
37	14 + 10 = 9 + __	0.92	0.84	6.30	6.97	1.43
38	29 + 3 = __ + 17	0.96	0.93	---	---	0.13

$X^2 = 87.94$ (38 items) X^2(items < 10) $= 70.43$ (37 items) $s^2 = 0.73$

Table 7

Predicted and observed proportions of errors and success-latency in fourth-grade addition, concept block 1, level 5

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
1	(20 + 1) + 8 = 24 + __	0.11	0.32	2.10	3.68	1.86
2	(12 + 0) + 14 = __ + 26	0.11	0.29	1.80	3.51	1.37
3	(4 + 16) + 8 = __ + 22	0.11	0.30	3.00	4.17	1.59
4	(23 + 0) + 5 = __ + 0	0.11	0.17	3.00	2.85	0.25
5	(2 + 16) + 8 = __ + 10	0.11	0.32	5.70	4.01	1.86
6	(11 + 0) + 11 = __ + 18	0.11	0.32	3.10	3.84	1.77
7	(14 + 3) + 4 = 20 + __	0.22	0.22	2.90	4.01	0.00
8	(12 + 7) + 9 = 27 + __	0.22	0.19	4.00	4.17	0.05
9	(14 + 12) + 2 = __ + 6	0.33	0.31	4.70	3.84	0.03
10	(26 + 0) + 0 = __ + 11	0.33	0.16	4.50	2.85	2.08
11	(18 + 4) + 7 = __ + 0	0.33	0.37	3.40	3.68	0.05
12	(10 + 2) + 9 = __ + 8	0.33	0.30	4.70	4.17	0.05
13	(0 + 16) + 6 = __ + 0	0.33	0.18	3.10	3.18	1.49
14	(8 + 9) + 8 = 2 + __	0.33	0.35	5.00	4.17	0.02
15	(15 + 6) + 4 = __ + 18	0.44	0.35	4.40	4.50	0.36
16	(10 + 0) + 18 = __ + 11	0.56	0.27	3.00	3.35	3.62
17	(17 + 10) + 1 = __ + 12	0.56	0.35	4.10	3.84	1.62
18	(14 + 2) + 11 = __ + 18	0.67	0.43	5.50	4.34	2.11
19	(14 + 11) + 2 = 18 + __	0.78	0.43	4.50	4.34	4.52

$\chi^2 = 24.69$ (19 items) χ^2(items < 10) = 24.69 (19 items) $s^2 = 1.17$

Table 8

Predicted and observed proportions of errors and success-latency in fourth-grade addition, concept block 3, level 2

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
1	32 + ___ = 33	0.07	0.15	2.50	3.08	0.33
2	35 + ___ = 37	0.07	0.16	2.80	3.07	0.39
3	37 + 0 = ___	0.07	0.12	3.20	3.23	0.14
4	39 + 0 = ___	0.07	0.13	2.80	3.35	0.20
5	33 + ___ = 36	0.07	0.16	4.00	3.18	0.43
6	34 + ___ = 35	0.07	0.15	2.10	3.03	0.34
7	32 + ___ = 32	0.07	0.14	2.70	3.02	0.29
8	47 + ___ = 47	0.07	0.15	1.80	2.66	0.36
9	41 + 7 = ___	0.07	0.26	4.90	4.44	1.34
10	43 + ___ = 43	0.07	0.15	1.90	2.76	0.34
11	42 + ___ = 42	0.07	0.15	1.00	2.78	0.33
12	47 + 0 = ___	0.07	0.18	3.20	3.85	0.55
13	45 + 0 = ___	0.08	0.16	3.00	3.72	0.29
14	43 + ___ = 44	0.08	0.16	2.90	2.82	0.24
15	45 + ___ = 46	0.08	0.16	2.80	2.77	0.25
16	33 + ___ = 34	0.14	0.15	2.40	3.05	0.00
17	38 + 2 = ___	0.14	0.39	4.90	5.50	1.82
18	39 + ___ = 40	0.14	0.28	2.40	4.04	0.65
19	32 + 3 = ___	0.14	0.18	4.30	3.98	0.07
20	44 + 1 = ___	0.14	0.29	2.50	4.77	0.71
21	47 + 3 = ___	0.14	0.50	3.60	6.03	3.66
22	43 + ___ = 46	0.14	0.17	2.30	2.94	0.04
23	___ + 1 = 47	0.14	0.49	4.70	6.02	3.36
24	___ + 0 = 44	0.17	0.29	2.90	4.79	0.42
25	45 + ___ = 49	0.17	0.18	3.50	2.96	0.01
26	41 + ___ = 45	0.17	0.18	3.90	3.05	0.00
27	___ + 0 = 47	0.17	0.32	3.50	4.97	0.63
28	49 + 1 = ___	0.17	0.43	3.80	5.64	1.71

Table 8 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
29	___ + 0 = 34	0.29	0.20	2.00	4.17	0.37
30	31 + 7 = ___	0.29	0.18	6.00	3.82	0.56
31	35 + ___ = 39	0.29	0.17	3.00	3.19	0.66
32	41 + 4 = ___	0.29	0.26	3.20	4.51	0.02
33	___ + 2 = 50	0.29	0.70	6.40	7.24	5.56
34	43 + 6 = ___	0.29	0.28	5.60	4.59	0.00
35	43 + 2 = ___	0.33	0.28	6.40	4.68	0.09
36	___ + 1 = 45	0.33	0.46	7.50	5.89	0.41
37	44 + 5 = ___	0.33	0.29	7.40	4.67	0.05
38	42 + 3 = ___	0.33	0.27	5.80	4.60	0.13
39	42 + 0 = ___	0.33	0.14	4.60	3.54	1.74
40	31 + ___ = 40	0.43	0.36	5.00	4.72	0.16
41	___ + 2 = 36	0.43	0.35	6.60	5.25	0.22
42	45 + 4 = ___	0.43	0.30	3.40	4.76	0.55
43	43 + ___ = 47	0.43	0.18	4.00	3.00	3.07
44	43 + 7 = ___	0.43	0.46	4.40	5.69	0.03
45	42 + ___ = 45	0.50	0.17	4.80	2.96	4.70
46	41 + 9 = ___	0.57	0.44	3.90	5.52	0.51
47	___ + 3 = 50	0.57	0.69	5.70	7.16	0.43
48	___ + 3 = 49	0.57	0.49	6.40	5.97	0.18
49	44 + ___ = 50	0.67	0.34	6.10	4.23	2.88
50	32 + ___ = 33	0.71	0.15	2.60	3.08	17.77
51	32 + 7 = ___	0.71	0.19	3.80	3.89	12.94
52	___ + 3 = 37	0.71	0.35	9.40	5.23	4.18
53	___ + 4 = 46	0.71	0.44	7.40	5.70	2.08
54	42 + 7 = ___	0.83	0.27	6.00	4.50	9.46
55	___ + 8 = 50	0.83	0.64	7.30	6.73	1.00
56	___ + 6 = 48	0.83	0.45	9.40	5.65	3.54
57	44 + 6 = ___	0.92	0.47	---	---	4.80

x^2 = 97.02 (57 items) x^2(items < 10) = 66.31 (55 items) s^2 = 2.29

Table 9

Predicted and observed proportions of errors and success-latency in fourth-grade additon, concept block 3, level 3

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
1	__ + 0 = 69	0.01	0.05	3.30	3.78	1.39
2	65 + 0 = __	0.01	0.02	2.20	2.29	0.25
3	__ + 0 = 62	0.01	0.07	3.60	3.72	1.92
4	66 + __ = 66	0.01	0.02	1.40	1.31	0.13
5	__ + 0 = 63	0.01	0.06	3.10	3.73	1.83
6	64 + __ = 65	0.01	0.03	1.90	2.15	0.64
7	__ + 0 = 70	0.01	0.05	2.90	3.78	1.32
8	39 + __ = 39	0.02	0.10	2.00	2.69	3.58
9	44 + __ = 47	0.02	0.11	2.90	3.19	1.95
10	51 + 4 = __	0.02	0.07	3.40	3.43	0.78
11	52 + __ = 52	0.02	0.05	1.60	2.03	0.28
12	61 + 4 = __	0.02	0.05	3.30	3.51	0.70
13	__ + 1 = 62	0.02	0.15	4.30	5.12	5.13
14	37 + __ = 37	0.04	0.11	2.10	2.79	2.68
15	34 + 1 = __	0.04	0.14	2.60	3.44	4.08
16	__ + 0 = 39	0.04	0.13	3.60	3.54	3.47
17	49 + __ = 49	0.04	0.06	2.20	2.18	0.12
18	__ + 0 = 48	0.04	0.10	3.90	3.61	0.94
19	48 + __ = 50	0.04	0.20	3.20	4.43	3.97
20	54 + __ = 57	0.05	0.06	4.00	2.68	0.09
21	54 + __ = 58	0.05	0.06	3.00	2.68	0.07
22	56 + 1 = __	0.05	0.07	3.10	3.62	0.26
23	51 + __ = 51	0.05	0.05	1.50	2.08	0.01
24	59 + __ = 59	0.05	0.03	1.80	1.67	0.15
25	__ + 0 = 60	0.05	0.07	4.00	3.70	0.20
26	51 + 5 = __	0.05	0.07	3.50	3.38	0.18
27	61 + __ = 65	0.05	0.04	2.50	3.33	0.13
28	67 + 3 = __	0.05	0.11	3.40	5.06	1.75

Table 9 (continued)

Rank	Equations	Observed $(1 - p_1)$	Predicted $(1 - p_1)$	Observed Latency	Predicted Latency	X^2
29	__ + 0 = 35	0.06	0.14	3.50	3.50	2.76
30	31 + 1 = __	0.06	0.15	2.40	3.42	3.16
31	61 + 7 = __	0.07	0.04	3.20	3.35	0.78
32	41 + 2 = __	0.08	0.11	2.60	3.45	0.64
33	__ + 1 = 33	0.08	0.30	4.10	4.88	12.31
34	31 + __ = 36	0.08	0.20	3.10	3.87	4.78
35	42 + __ = 44	0.08	0.12	2.30	3.28	0.41
36	43 + 4 = __	0.08	0.14	4.90	4.09	0.70
37	52 + 2 = __	0.09	0.08	2.90	3.54	0.05
38	__ + 1 = 67	0.10	0.13	4.10	5.16	0.40
39	65 + 3 = __	0.10	0.05	3.50	3.59	1.96
40	46 + 0 = __	0.11	0.04	2.90	2.14	6.61
41	31 + 2 = __	0.12	0.14	2.60	3.37	0.20
42	43 + __ = 45	0.12	0.12	3.30	3.23	0.01
43	48 + __ = 49	0.12	0.09	2.70	2.97	0.80
44	63 + 7 = __	0.12	0.10	4.50	4.83	0.17
45	__ + 1 = 69	0.12	0.12	4.60	5.17	0.00
46	__ + 0 = 49	0.14	0.10	3.60	3.62	1.29
47	54 + 4 = __	0.14	0.07	3.40	3.45	1.79
48	51 + __ = 57	0.14	0.07	3.30	2.85	1.81
49	36 + 3 = __	0.14	0.12	3.20	3.36	0.23
50	32 + __ = 36	0.14	0.20	3.50	3.81	1.12
51	43 + 3 = __	0.15	0.09	3.40	3.41	2.50
52	31 + __ = 37	0.16	0.20	3.20	3.87	0.49
53	42 + __ = 48	0.16	0.11	4.40	3.31	0.67
54	43 + __ = 48	0.16	0.11	4.30	3.25	0.77
55	45 + 3 = __	0.16	0.09	3.90	3.43	1.51
56	46 + __ = 46	0.17	0.07	2.20	2.33	10.76
57	__ + 4 = 66	0.17	0.12	5.00	4.97	0.81
58	__ + 2 = 40	0.18	0.47	5.60	6.34	17.72
59	41 + 0 = __	0.18	0.05	2.80	2.10	24.98

Table 9 (continued)

Rank	Equations	Observed $(1 - p_1)$	Predicted $(1 - p_1)$	Observed Latency	Predicted Latency	x^2
60	51 + 7 = ___	0.18	0.06	3.40	3.27	5.83
61	57 + ___ = 60	0.18	0.08	4.40	3.25	3.27
62	54 + ___ = 60	0.18	0.09	4.70	3.43	2.62
63	34 + ___ = 38	0.20	0.18	2.90	3.70	0.13
64	41 + 1 = ___	0.20	0.11	2.70	3.50	4.75
65	41 + 6 = ___	0.20	0.08	3.60	3.25	10.73
66	41 + ___ = 48	0.20	0.11	4.00	3.37	1.94
67	___ + 0 = 43	0.21	0.11	3.30	3.57	6.58
68	46 + 1 = ___	0.23	0.10	2.40	3.54	12.52
69	___ + 2 = 55	0.23	0.18	6.20	5.00	0.40
70	42 + 4 = ___	0.24	0.09	4.40	3.36	6.51
71	45 + 5 = ___	0.24	0.18	3.80	4.78	1.43
72	47 + ___ = 50	0.24	0.14	2.70	3.76	6.14
73	___ + 2 = 38	0.26	0.27	5.70	4.87	0.03
74	44 + ___ = 44	0.26	0.07	1.50	2.43	32.14
75	___ + 3 = 64	0.27	0.14	5.30	5.01	6.19
76	52 + 8 = ___	0.27	0.13	5.60	4.69	3.86
77	___ + 5 = 39	0.28	0.24	4.90	4.70	0.30
78	___ + 4 = 48	0.28	0.20	6.00	4.83	1.02
79	___ + 2 = 43	0.28	0.24	6.50	4.91	0.26
80	___ + 5 = 50	0.28	0.37	6.00	6.24	0.90
81	41 + 8 = ___	0.28	0.08	4.60	3.14	14.97
82	42 + ___ = 47	0.30	0.11	3.20	3.30	24.15
83	___ + 7 = 38	0.31	0.24	4.40	4.57	1.70
84	___ + 3 = 69	0.32	0.12	5.50	5.05	15.62
85	___ + 2 = 47	0.32	0.21	7.10	4.94	1.64
86	___ + 4 = 50	0.33	0.38	6.30	6.30	0.60
87	___ + 5 = 46	0.33	0.20	5.10	4.75	6.70
88	___ + 2 = 60	0.36	0.32	5.60	6.50	0.17
89	___ + 9 = 50	0.38	0.34	5.40	6.00	0.38
90	___ + 6 = 38	0.39	0.24	6.00	4.63	6.24

Table 9 (continued)

Rank	Equations	Observed $(1 - p_1)$	Predicted $(1 - p_1)$	Observed Latency	Predicted Latency	X^2
91	___ + 3 = 45	0.44	0.22	5.50	4.86	7.04
92	___ + 3 = 40	0.47	0.46	6.00	6.28	0.02
93	___ + 5 = 49	0.52	0.19	5.70	4.78	45.44
94	___ + 4 = 45	0.52	0.21	6.90	4.80	13.81
95	___ + 4 = 60	0.55	0.31	7.00	6.38	5.74

$X^2 = 391.87$ (95 items) $X^2(\text{items} < 10) = 156.75$ (83 items) $s^2 = 0.62$

Table 10

Predicted and observed proportions of errors and success-latency in fourth-grade addition, concept block 3, level 4

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
1	2 + 46 = 0 + ___	0.02	0.17	3.70	4.14	7.23
2	7 + 50 = 0 + ___	0.02	0.12	3.20	3.80	3.68
3	0 + 46 = 0 + ___	0.06	0.05	2.50	2.70	0.28
4	1 + 54 = ___ + 0	0.07	0.17	3.00	4.24	3.24
5	56 + 1 = 54 + ___	0.11	0.45	3.50	5.70	20.61
6	69 + 0 = ___ + 9	0.13	0.12	3.20	3.98	0.12
7	47 + 0 = 45 + ___	0.15	0.16	2.90	4.12	0.12
8	48 + 21 = 21 + ___	0.17	0.66	5.90	6.77	24.63
9	32 + 20 = 0 + ___	0.21	0.17	5.20	4.20	0.44
10	4 + 50 = 3 + ___	0.23	0.25	4.50	4.76	0.16
11	63 + 4 = 7 + ___	0.26	0.37	5.30	5.44	1.13
12	57 + 0 = ___ + 7	0.27	0.12	3.30	3.80	5.36
13	7 + 40 = 6 + ___	0.29	0.35	3.90	5.19	0.79
14	4 + 42 = 1 + ___	0.29	0.33	4.90	5.07	0.37
15	7 + 54 = 1 + ___	0.30	0.57	5.60	6.23	6.85
16	52 + 10 = ___ + 11	0.30	0.50	5.30	5.96	3.50
17	45 + 12 = ___ + 12	0.36	0.61	4.10	6.40	5.82
18	8 + 51 = 5 + ___	0.36	0.36	4.60	5.34	0.00
19	23 + 30 = ___ + 2	0.36	0.34	6.20	5.19	0.04
20	5 + 52 = 3 + ___	0.36	0.35	4.20	5.27	0.01
21	22 + 20 = ___ + 10	0.40	0.36	5.40	5.20	0.22
22	9 + 42 = 1 + ___	0.41	0.56	5.20	6.08	2.09
23	30 + 26 = 10 + ___	0.41	0.38	5.20	5.40	0.17
24	4 + 64 = ___ + 2	0.44	0.36	5.30	5.42	0.54
25	12 + 56 = ___ + 0	0.44	0.26	6.30	4.90	3.84
26	67 + 1 = ___ + 30	0.44	0.36	7.00	5.39	0.61
27	7 + 37 = ___ + 2	0.44	0.56	7.10	5.99	2.80
28	7 + 45 = 3 + ___	0.46	0.77	5.10	7.07	12.31

Table 10 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
29	11 + 40 = 3 + __	0.46	0.57	5.70	6.12	1.20
30	47 + 8 = 1 + __	0.46	0.57	6.80	6.14	2.24
31	50 + 19 = 45 + __	0.52	0.55	6.60	6.26	0.05
32	36 + 33 = 13 + __	0.52	0.63	6.90	6.60	1.20
33	24 + 22 = __ + 20	0.54	0.52	7.40	5.93	0.08
34	32 + 22 = 20 + __	0.55	0.53	5.40	6.05	0.02
35	60 + 5 = 21 + __	0.57	0.37	6.90	5.43	3.77
36	5 + 36 = __ + 4	0.58	0.67	6.60	6.46	1.67
37	29 + 30 = 13 + __	0.59	0.51	5.60	5.98	0.60
38	10 + 41 = __ + 7	0.59	0.59	5.50	6.20	0.00
39	22 + 31 = __ + 20	0.59	0.53	6.90	6.04	0.64
40	51 + 13 = __ + 41	0.61	0.63	6.80	6.52	0.03
41	33 + 35 = __ + 13	0.61	0.63	6.50	5.68	0.05
42	48 + 10 = 23 + __	0.64	0.49	5.70	5.90	1.77
43	15 + 40 = __ + 7	0.64	0.59	6.10	6.26	0.18
44	54 + 4 = __ + 27	0.64	0.47	6.10	5.78	4.94
45	5 + 56 = __ + 3	0.65	0.58	7.20	6.27	0.46
46	50 + 2 = __ + 13	0.68	0.45	5.30	5.65	4.66
47	23 + 29 = 8 + __	0.68	0.85	4.60	7.65	5.07
48	42 + 2 = __ + 9	0.69	0.56	7.60	5.99	3.30
49	42 + 15 = 32 + __	0.71	0.63	6.70	6.46	1.16
50	29 + 26 = __ + 5	0.71	0.69	6.60	6.69	0.05
51	8 + 36 = 3 + __	0.71	0.56	6.40	6.02	4.16
52	51 + 8 = __ + 35	0.73	0.49	6.90	5.88	10.15
53	37 + 15 = 31 + __	0.73	0.80	7.10	7.32	1.66
54	19 + 26 = __ + 3	0.73	0.67	7.60	6.50	0.73
55	23 + 25 = __ + 14	0.75	0.61	7.60	6.31	3.85
56	17 + 37 = 13 + __	0.75	0.80	6.10	7.31	0.69
57	54 + 4 = 23 + __	0.77	0.47	6.80	5.78	8.18
58	29 + 39 = __ + 21	0.78	0.83	8.10	7.69	0.38

Table 10 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
59	21 + 47 = __ + 16	0.78	0.64	7.40	6.65	1.97
60	39 + 29 = 13 + __	0.78	0.81	7.40	7.52	0.12
61	25 + 39 = __ + 13	0.78	0.81	7.80	7.46	0.09
62	17 + 24 = 14 + __	0.81	0.91	7.00	8.08	4.81
63	24 + 27 = __ + 16	0.82	0.91	8.50	8.27	2.45
64	36 + 17 = 8 + __	0.82	0.85	8.00	7.66	0.21
65	16 + 38 = __ + 13	0.82	0.80	7.30	7.31	0.09
66	25 + 19 = __ + 17	0.83	0.91	6.90	8.18	3.61
67	29 + 12 = __ + 5	0.83	0.84	7.60	7.42	0.01
68	28 + 25 = 13 + __	0.84	0.80	7.30	7.30	0.47
69	52 + 1 = 24 + __	0.84	0.67	8.50	6.58	5.68
70	28 + 18 = __ + 21	0.85	0.81	7.80	7.30	0.67
71	31 + 24 = __ + 19	0.89	0.82	6.60	7.45	1.41
72	19 + 22 = __ + 4	0.90	0.84	7.90	7.40	1.25
73	35 + 20 = 28 + __	0.91	0.74	7.80	7.01	3.23
74	53 + 1 = __ + 16	0.91	0.67	7.60	6.59	11.10
75	35 + 9 = 23 + __	0.92	0.69	9.00	6.61	11.37
76	11 + 51 = 9 + __	0.96	0.71	7.10	6.88	6.79

$x^2 = 225.08$ (76 items) x^2(items < 10) $= 134.90$ (70 items) $s^2 = 1.04$

Table 11

Predicted and observed proportions of errors and success-latency in fifth-grade addition, concept block 3, levels 3 and 4

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
1	__ + 0 = 34	0.14	0.07	2.90	1.29	0.49
2	32 + __ = 33	0.14	0.13	1.60	2.67	0.01
3	35 + __ = 37	0.14	0.18	1.80	3.10	0.05
4	31 + __ = 40	0.14	0.36	3.70	4.32	1.18
5	37 + 0 = __	0.14	0.07	1.50	1.09	0.47
6	39 + __ = 40	0.14	0.25	0.60	3.77	0.34
7	39 + 0 = __	0.14	0.08	2.00	1.27	0.29
8	32 + 3 = __	0.14	0.11	2.50	2.05	0.06
9	32 + __ = 38	0.14	0.23	2.60	3.47	0.28
10	33 + __ = 36	0.14	0.18	1.20	3.08	0.05
11	34 + __ = 35	0.14	0.15	1.40	2.85	0.00
12	35 + __ = 39	0.14	0.23	1.40	3.42	0.24
13	32 + __ = 32	0.14	0.11	2.20	2.51	0.05
14	7 + 29 = 34 + __	0.14	0.31	4.70	4.74	1.75
15	32 + 6 = __ + 36	0.14	0.27	2.60	4.13	1.17
16	34 + 5 = 35 + __	0.14	0.32	2.40	4.35	1.93
17	33 + 4 = __ + 0	0.14	0.16	3.90	2.96	0.03
18	33 + __ = 34	0.17	0.14	0.90	2.76	0.04
19	38 + 2 = __	0.17	0.22	3.30	3.37	0.09
20	32 + 7 = __	0.17	0.18	3.90	2.69	0.01
21	31 + 7 = __	0.17	0.17	3.20	2.60	0.00
22	__ + 3 = 37	0.17	0.16	5.10	2.70	0.01
23	31 + 17 = 44 + __	0.20	0.54	4.90	5.63	7.16
24	15 + 22 = __ + 34	0.21	0.32	5.00	4.57	0.68
25	5 + 41 = 40 + __	0.27	0.41	2.80	4.58	1.22
26	2 + 46 = __ + 43	0.27	0.51	3.80	5.49	3.59
27	4 + 35 = 4 + __	0.29	0.32	3.30	4.35	0.06
28	1 + 37 = 24 + __	0.29	0.31	7.20	4.52	0.03

Table 11 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
29	__ + 2 = 36	0.33	0.14	4.60	2.54	1.92
30	21 + 25 = 43 + __	0.33	0.54	2.90	5.85	2.71
31	18 + 19 = 10 + __	0.36	0.54	6.40	5.99	1.81
32	7 + 30 = __ + 28	0.36	0.46	5.70	5.45	0.63
33	21 + 15 = 15 + __	0.36	0.54	3.10	5.78	1.91
34	24 + 12 = __ + 4	0.43	0.31	7.10	4.55	0.88
35	22 + 16 = __ + 10	0.43	0.44	6.20	5.14	0.01
36	20 + 19 = 5 + __	0.43	0.38	4.20	4.89	0.13
37	26 + 11 = 26 + __	0.43	0.50	3.80	5.59	0.25
38	11 + 26 = __ + 7	0.43	0.38	5.60	4.85	0.17
39	28 + 21 = 11 + __	0.47	0.72	7.20	6.67	4.77
40	11 + 28 = __ + 3	0.47	0.35	5.30	4.75	0.87
41	3 + 36 = __ + 14	0.50	0.35	7.70	4.75	1.35
42	47 + 2 = __ + 13	0.53	0.53	5.80	5.58	0.00
43	26 + 23 = 6 + __	0.53	0.60	5.10	5.85	0.25
44	9 + 29 = 12 + __	0.57	0.54	6.00	6.01	0.05
45	20 + 28 = __ + 18	0.60	0.75	6.70	6.59	1.78
46	10 + 39 = 7 + __	0.67	0.56	6.20	5.45	0.75
47	7 + 41 = 13 + __	0.67	0.59	7.00	5.83	0.35
48	23 + 23 = 13 + __	0.67	0.70	5.30	6.53	0.07
49	0 + 48 = __ + 12	0.67	0.42	6.70	4.88	3.69
50	0 + 47 = 12 + __	0.67	0.40	5.70	4.79	4.36
51	39 + 7 = 12 + __	0.67	0.66	7.30	6.59	0.00
52	14 + 24 = __ + 16	0.71	0.56	6.80	5.89	1.27
53	39 + 7 = 9 + __	0.73	0.76	8.30	7.52	0.04
54	27 + 11 = 25 + __	0.79	0.50	7.70	5.61	4.60
55	12 + 34 = __ + 28	0.93	0.77	9.40	7.40	2.18
56	19 + 27 = 38 + __	0.93	0.72	6.60	7.13	3.27
57	34 + 13 = 29 + __	0.97	0.80	---	---	2.66

X^2 = 64.00 (57 items) X^2(items < 10) = 64.00 (57 items) S^2 = 2.32

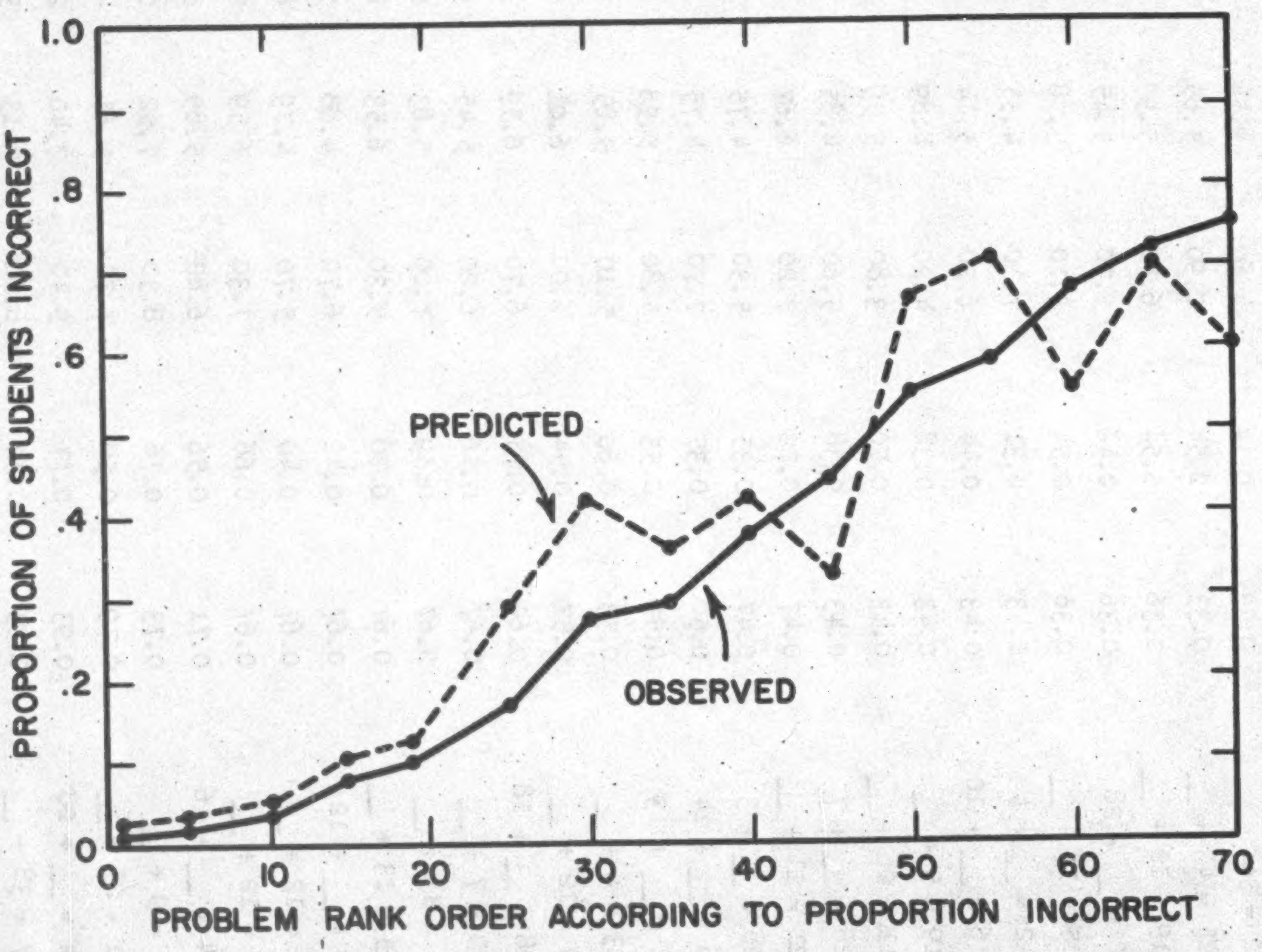

Figure 3. Predicted and observed proportions of errors on fourth-grade addition.

Some of the X^2 values obtained, as for example in Table 4, are extremely high and would usually be an indication of a poor fit, but a closer look at the components of the X^2 shows that 5 of the 38 problems in this analysis contribute more than two thirds of the total X^2. In the particular case cited, the large reduction in X^2 still does not make the value of X^2 such that the model would not normally be rejected. When we do reduce the X^2 values shown in Tables 4-11 by removing the few extreme components whose individual contributions are equal to or greater than 10, we find that in four of the eight cases we obtain a X^2 value whose probability P is such that $.1 < P < .9$ under the null hypothesis. Since calculation of the regression coefficients included the extreme problems, a recalculation of the regression coefficients omitting the few extreme problems from the data would yield better fits of the model to data than those obtained.

The overall X^2's as well as the reduced X^2's are shown at the bottom of Tables 4-11. Perusal of these tables with particular attention paid to the items for which predictions are unsatisfactory, and thus the X^2 contribution high, suggests immediately further analysis that incorporates variables designed to handle special algorithms. Moreover, in those cases for which $P < .1$, it should be noted that the actual predictions are mostly fairly good. Our viewpoint on this matter is that we hardly expected to fit the data exactly with such a small number of variables.

Figure 3 presents a graph of the predicted and observed proportions

---

Insert Figure 3 about here

---

of errors as a function of observed rank order of observed difficulty. The data for these curves were drawn from fourth-grade addition, block 1, levels 2, 3 and 4. An inspection of the two curves shows a relatively good fit for the regression model, even in the heterogeneous case of problems drawn from different drills, levels and correspondingly different groups of children.

There are several qualitative observations about the data of Tables 4-11 we want to make at this point. In the first place, the first problem presented on each day was deleted from all analyses when the results showed a short but significant warm-up effect. This rendered the initial problem more difficult independent of structural variables. Happily it was not necessary to include order of presentation as a variable since there was no significant warm-up effect beyond the first problem of a drill. The sequential effects, if any, of errors on the immediately following problem have not yet been analyzed systematically, but again this does not appear to be a very strong effect. Although we intend to go into this question in more detail on a subsequent occasion, the assumption of statistical independence of problem-items seems to be correct to a first approximation.

Tables 4, 5, 8, and 9 report data on problems of the general form m + n = p, where any of m, n or p may be two-digit numbers. What is striking is that the hardest problems are to a very large extent of the form __ + n = p. The last 2 problems of Table 4 are of this kind (the problem being ranked in order of difficulty from easiest to hardest), the last 7 of Table 5, and the last 13 of Table 9. The effect is not as noticeable in Table 8, although over half of the last 11 problems are of this form. Moreover, if we look at the easiest problems in these same tables, the form $\_\_ + n = p$ is very much excluded. With the exception of $\_\_ + 0 = p$, it does not occur in the easy half of Table 4, in Table 5 the form does not appear among the first 19 least difficult items, and in Table 8, not among the first 22. The evidence on this point is more mixed in Table 9. All in all, these results suggest that the transformation steps defined in the theoretical section might well be broken into separately weighted classes to differentiate $\_\_ + n = p$ from $m + \_\_ = p$. In some preliminary efforts aimed at refining and improving the predictive results reported here we have had some success with this distinction.

Although the predictive results from Tables 6, 7, 10 and 11 are far from the best that a mature theory should be able to offer, we are not dissatisfied with them as a beginning because of the relative difficulty of intuitively rank ordering the expected error rate of problems of the form $ab + cd = ef + gh$. The three variables that we consider bring a surprising amount of order to what appears at first glance to be a quite complex set of problem-items.

We turn now to the success-latency data for the same problems of fourth- and fifth-grade addition. The predicted and observed latencies are also given in Tables 4-11, with the predicted values depending on the appropriate regression coefficients of Table 3. As is clear from Table 3, the multiple correlations obtained for the fit of the predicted latencies are very comparable to those obtained for the predicted responses, and indicate that the success-latency data are as regular in range of variation as the response data.

In the analysis of latencies we have restricted ourselves to the success-latencies, i.e., the latencies of correct responses, because of their direct relevance for the analysis of the structure of the algorithms students use. Although error-latencies also contain much useful information, they include latencies of random guesses, false starts and other heterogeneous factors that are not easily disentangled. In a few cases latency data were garbled in transmission from the school to the computer, and in such cases we have simply entered a blank in both the predicted and observed columns.

There are various ways of evaluating the overall fit of the latency predictions reported in these tables. The statistic S^2, already mentioned, is given at the bottom of each table. Although this statistic may be used to find a significance level for the fit of the structural models, at this stage of investigation it seems more useful to interpret S^2 directly in terms of the quantitative closeness of the predictions. When the structural variables f_{ij} are not random variables, then S^2 is a good estimator of the variance σ^2 of the errors in the prediction of the models. Taking the algebraic sign into account, the expectation of these errors is nearly zero and the assumption that they are normally distributed with variance σ^2 is approximately satisfied also, and so we may evaluate the predictions of each table by looking at the magnitude of S, the bulk of the errors being within one standard deviation of the observed values. The values of S for Tables 4-11 are .81, 1.14, .85, 1.08, 1.51, .79, 1.02 and 1.52, respectively, which may be interpreted to mean that errors of prediction greater than 1 or 1.5 seconds do not occur very often. From inspection of the tables it may also be seen that the observed values have a range from about 3 seconds (Table 4) to more than 8.5 seconds (Table 11), and consequently, predictions of this accuracy are far from perfect, yet good enough to be practically useful.

Still another useful measure is the average percentage error of the predictions. If there are n items in a table, if o_i is the mean observed success-latency for item i, and p_i is the predicted latency, then the average percentage error (A.E.) is defined by

$$A.E = \frac{100}{n} \sum_{i=1}^{n} \frac{|o_i - p_i|}{p_i} .$$

This measure for Tables 4-11 has the values 16.4%, 19.8%, 12.3%, 20.6%, 25.7%, 16.6%, 12.8% and 31.4%, respectively.

The qualitative remarks made about proportion of errors pretty well apply to success-latencies, as would be expected because of the high positive correlation between the two variables. This is particularly true of latencies for problems of the form __ + n = p, as the reader may easily conclude from an inspection of Tables 4-11.

Figure 4 presents a graph of the predicted and observed success-latencies for the same problem-items for which response predictions are shown in Figure 3. The predicted and observed latencies are plotted as a function of the rank of observed latency, and consequently, the curve of observed latencies is monotonically increasing and smoother than the predicted curve, but the fit is qualitatively reasonably good.

Insert Figure 4 about here

Subtraction--grades four and five. The three independent variables used in the linear regression analyses of subtraction were NSTEPS as described previously and the two magnitude variables, magnitude of the difference (MAGDIF) and magnitude of the subtrahend (MAGSUB). The values of MAGDIFF and MAGSUB are not affected by the problem format. For example, in all three problems, 31 - 16 = __, 31 - __ = 15 and __ - 16 = 15, MAGDIF has the value 15 and MAGSUB the value 16.

The coefficients obtained for the regression equations are shown in Table 12, which is laid out in a manner identical to that of Table 3.

Insert Table 12 about here

As in the case of Table 3 it is clear that NSTEPS is the most important of the three variables in predicting errors or success-latencies. Also the values obtained are comparable to those given in Table 3. In the confines of the present paper it has not been possible to explore the possibility of a joint analysis of addition and subtraction, with a particular emphasis on process variables like NSTEPS, but this is a clearly indicated direction for future research.

Again, as in the case of Table 3, the multiple correlation coefficients shown in Table 12 indicate that the three independent variables are accounting for a good deal of the variation in the observed response proportions and success-latencies.

Tables 13-19 present predicted and observed proportions of errors, and other information identical to that given in Tables 4-11 for addition.

Insert Tables 13, 14, 15, 16, 17, 18 and 19 about here

The overall X^2's for subtraction exhibit a pattern very similar to those

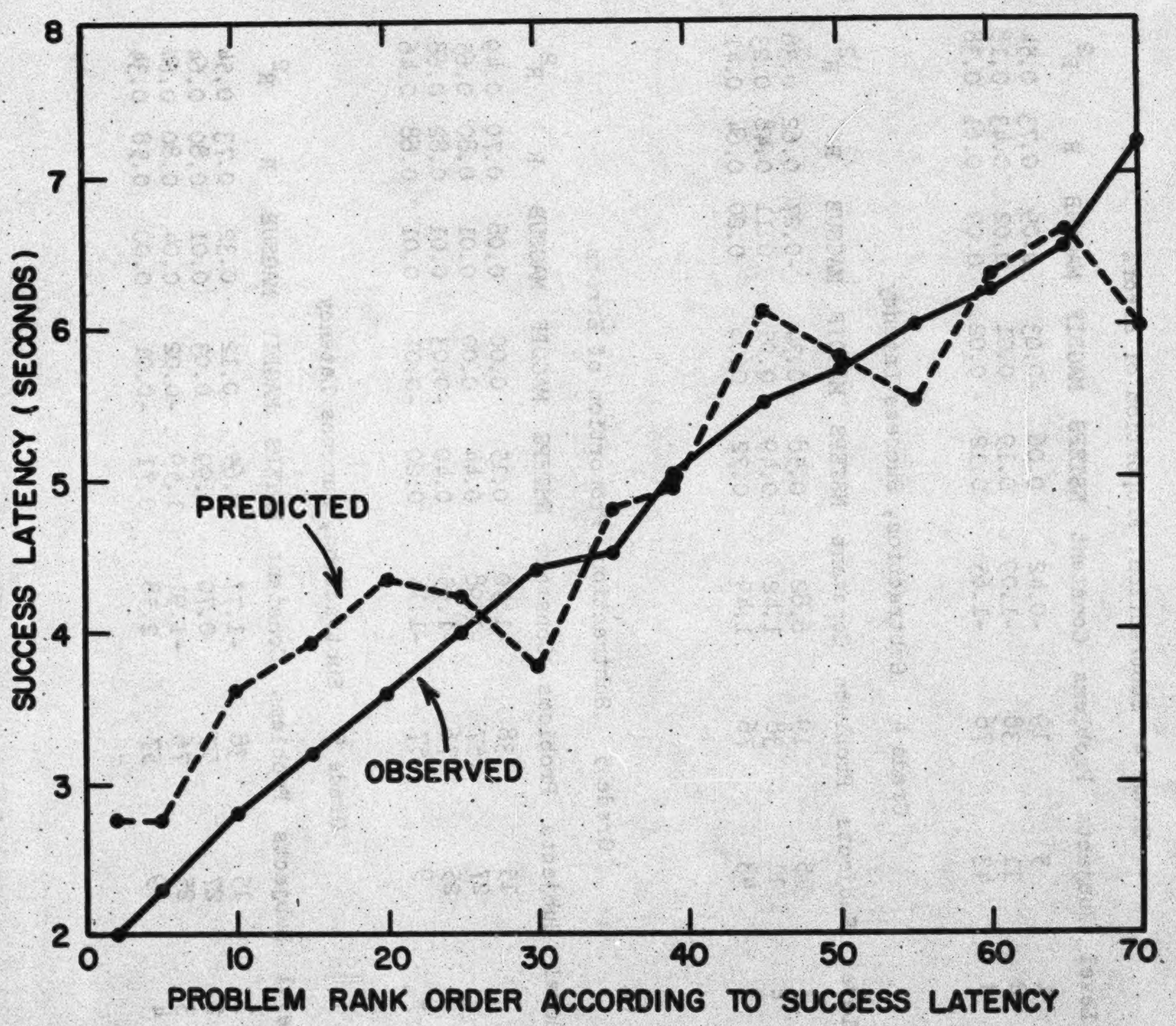

Figure 4. Predicted and observed success-latencies for fourth-grade addition.

Table 12

Regression Coefficients for Subtraction

Grade 4 Subtraction, Proportion of Errors

Level	Subjects	Problems	Constant	NSTEPS	MAGDIF	MAGSUB	R	R^2
1	5	19	-0.42	0.06	-0.03	0.09	0.73	0.54
2	11	38	-1.09	0.19	0.01	0.02	0.43	0.18
3	43	76	-1.63	0.12	0.02	0.09	0.61	0.38

Grade 4 Subtraction, Success Latency

Level	Subjects	Problems	Constant	NSTEPS	MAGDIF	MAGSUB	R	R^2
1	5	19	6.82	0.58	-0.34	-0.27	0.62	0.38
2	11	38	1.42	0.49	0.05	0.11	0.48	0.23
3	43	76	1.49	0.32	0.06	0.20	0.64	0.41

Grade 5 Subtraction, Proportion of Errors

Level	Subjects	Problems	Constant	NSTEPS	MAGDIF	MAGSUB	R	R^2
1	15	38	-1.50	0.15	0.00	0.08	0.70	0.49
2	27	57	-1.98	0.44	0.00	0.01	0.80	0.65
3	25	76	-1.65	0.40	-0.03	0.01	0.82	0.68
4	9	57	-1.14	0.20	-0.01	0.01	0.68	0.46

Grade 5 Subtraction, Success Latency

Level	Subjects	Problems	Constant	NSTEPS	MAGDIF	MAGSUB	R	R^2
1	15	38	-1.77	0.65	0.12	0.32	0.73	0.54
2	27	57	0.70	0.99	0.03	0.01	0.80	0.64
3	25	76	-1.91	1.59	-0.02	0.04	0.80	0.64
4	9	57	2.58	0.71	-0.01	0.00	0.58	0.34

Table 13

Predicted and observed proportions of errors and success-latency in fourth-grade subtraction, level 1

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
1	12 - 0 = __	0.10	0.16	2.70	2.72	0.13
2	14 - 0 = __	0.10	0.14	1.70	2.04	0.08
3	__ - 3 = 11	0.20	0.38	7.70	4.58	0.69
4	12 - 6 = __	0.20	0.60	3.00	5.48	3.33
5	14 - 3 = __	0.40	0.32	4.70	3.42	0.14
6	__ - 1 = 13	0.40	0.29	4.90	5.03	0.27
7	__ - 7 = 6	0.40	0.70	8.80	6.37	2.21
8	__ - 0 = 13	0.40	0.19	2.50	3.55	1.48
9	13 - __ = 7	0.40	0.62	7.00	5.72	1.00
10	14 - __ = 5	0.40	0.77	6.40	5.59	3.77
11	__ - 6 = 5	0.60	0.67	7.80	6.99	0.12
12	12 - __ = 8	0.60	0.50	5.00	5.92	0.18
13	11 - 4 = __	0.60	0.49	6.80	5.68	0.26
14	14 - __ = 4	0.80	0.77	4.00	4.49	0.03
15	15 - __ = 8	0.80	0.65	3.20	5.11	0.50
16	__ - 5 = 8	0.80	0.59	2.70	6.23	0.94
17	__ - 10 = 2	0.90	0.79	---	---	0.38
18	14 - __ = 6	0.90	0.72	---	---	0.82
19	11 - __ = 2	0.90	0.80	---	---	0.33

$\chi^2 = 16.65$ (19 items) χ^2(items < 10) = 16.65 (19 items) $s^2 = 3.84$

Table 14

Predicted and observed proportions of errors and success-latency in fourth-grade subtraction, level 2

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
1	15 - 5 = ___	0.05	0.22	2.70	3.48	1.98
2	17 - 7 = ___	0.05	0.24	2.60	3.69	2.25
3	12 - ___ = 11	0.05	0.36	2.60	4.09	4.77
4	12 - 10 = ___	0.05	0.24	3.20	3.58	2.28
5	___ - 0 = 19	0.09	0.21	4.00	3.43	0.90
6	___ - 0 = 13	0.09	0.19	3.10	3.11	0.71
7	16 - ___ = 16	0.18	0.37	2.60	4.25	1.66
8	19 - ___ = 10	0.18	0.34	4.30	4.39	1.28
9	11 - 10 = ___	0.18	0.24	1.60	3.52	0.18
10	15 - ___ = 6	0.27	0.54	5.20	5.16	3.09
11	16 - 10 = ___	0.27	0.25	3.70	3.79	0.03
12	20 - ___ = 19	0.27	0.50	3.00	5.01	2.21
13	17 - ___ = 16	0.27	0.38	2.60	4.36	0.54
14	15 - 10 = ___	0.27	0.25	2.00	3.74	0.04
15	___ - 3 = 9	0.27	0.59	6.90	5.17	4.46
16	20 - ___ = 16	0.27	0.62	4.60	5.66	5.80
17	20 - ___ = 13	0.36	0.65	5.90	5.82	3.80
18	19 - ___ = 13	0.36	0.42	6.30	4.73	0.16
19	19 - ___ = 11	0.36	0.44	6.20	4.83	0.25
20	___ - 1 = 13	0.36	0.37	4.40	4.19	0.00
21	11 - 7 = ___	0.46	0.40	4.00	4.35	0.14
22	15 - ___ = 8	0.46	0.52	4.20	5.05	0.20
23	17 - 8 = ___	0.55	0.43	4.20	4.72	0.59
24	18 - ___ = 13	0.55	0.41	4.80	4.62	0.80
25	13 - 2 = ___	0.55	0.20	4.30	3.21	8.14
26	___ - 9 = 10	0.55	0.34	5.40	4.39	1.98
27	14 - 3 = ___	0.55	0.21	6.40	3.32	7.58

Table 14 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
28	__ - 1 = 16	0.64	0.38	4.10	4.36	3.06
29	__ - 3 = 14	0.64	0.39	5.60	4.46	2.69
30	__ - 5 = 6	0.64	0.60	4.50	5.22	0.07
31	19 - 6 = __	0.64	0.24	4.90	3.75	9.64
32	__ - 3 = 15	0.64	0.40	4.10	4.52	2.60
33	__ - 4 = 15	0.64	0.41	5.00	4.62	2.34
34	15 - __ = 9	0.73	0.51	3.50	5.00	1.99
35	20 - 5 = __	0.73	0.42	4.80	4.73	4.24
36	11 - 3 = __	0.73	0.37	4.40	4.14	5.95
37	__ - 5 = 11	0.91	0.40	6.80	4.51	11.60
38	__ - 7 = 4	0.91	0.61	7.80	5.33	4.10

$x^2 = 104.03$ (38 items) x^2(items < 10) $= 92.43$ (37 items) $s^2 = 1.80$

Table 15

Predicted and observed proportions of errors and success-latency in fourth-grade subtraction, level 3

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
1	19 - __ = 19	0.01	0.13	2.00	3.93	6.65
2	15 - __ = 15	0.01	0.11	1.70	3.68	5.53
3	13 - __ = 13	0.01	0.10	1.70	3.56	5.04
4	8 - 4 = __	0.01	0.08	1.90	2.84	4.47
5	8 - 8 = __	0.01	0.11	1.30	3.06	7.62
6	11 - 9 = __	0.02	0.33	2.10	4.64	22.67
7	26 - __ = 26	0.02	0.16	2.50	4.36	3.90
8	9 - 9 = __	0.03	0.13	1.30	3.26	7.95
9	8 - 2 = __	0.04	0.06	2.60	2.57	0.46
10	17 - 0 = __	0.04	0.04	2.30	2.54	0.05
11	__ - 1 = 21	0.04	0.16	4.50	4.25	2.97
12	9 - 7 = __	0.05	0.12	2.30	3.24	3.59
13	9 - __ = 6	0.05	0.09	2.90	3.08	1.26
14	__ - 0 = 2	0.05	0.04	3.30	2.24	0.19
15	__ - 0 = 3	0.05	0.04	2.70	2.31	0.13
16	10 - 6 = __	0.05	0.22	2.00	4.18	12.95
17	3 - __ = 0	0.06	0.05	2.50	2.39	0.13
18	__ - 1 = 0	0.06	0.05	3.20	2.32	0.48
19	12 - 0 = __	0.07	0.04	1.80	2.23	0.40
20	12 - __ = 6	0.08	0.28	2.10	4.62	11.32
21	3 - __ = 0	0.08	0.05	2.50	2.39	0.76
22	4 - __ = 4	0.08	0.03	2.70	2.05	4.10
23	25 - __ = 23	0.08	0.21	3.60	4.57	2.65
24	10 - __ = 7	0.09	0.18	2.70	4.09	4.44
25	19 - 10 = __	0.09	0.32	2.30	4.64	12.07
26	5 - __ = 3	0.10	0.06	2.80	2.70	1.84
27	11 - 4 = __	0.11	0.17	3.40	3.97	1.28
28	__ - 4 = 0	0.11	0.08	3.20	2.90	0.93

Table 15 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
29	19 - __ = 9	0.15	0.38	3.20	4.96	11.46
30	__ - 4 = 10	0.17	0.15	5.20	3.84	0.14
31	11 - __ = 4	0.19	0.31	4.50	4.69	3.63
32	__ - 2 = 13	0.19	0.15	5.10	3.95	0.71
33	18 - 2 = __	0.20	0.10	4.40	3.51	1.53
34	13 - 3 = __	0.20	0.10	4.90	3.33	1.67
35	__ - 1 = 19	0.21	0.19	5.20	4.44	0.12
36	12 - 4 = __	0.21	0.18	3.90	4.03	0.34
37	29 - __ = 27	0.23	0.23	3.20	4.82	0.00
38	__ - 2 = 19	0.23	0.27	5.40	4.95	0.24
39	__ - 1 = 8	0.24	0.08	4.00	3.13	27.17
40	14 - __ = 5	0.26	0.42	5.10	5.14	5.01
41	__ - 3 = 6	0.27	0.11	4.10	3.40	19.64
42	__ - 2 = 4	0.27	0.08	4.10	3.08	33.51
43	11 - __ = 5	0.27	0.18	3.80	3.92	0.73
44	26 - __ = 22	0.27	0.27	4.10	4.90	0.00
45	20 - 3 = __	0.28	0.20	4.00	4.40	2.23
46	__ - 10 = 2	0.30	0.37	5.20	4.84	1.14
47	26 - __ = 21	0.31	0.31	4.40	5.04	0.00
48	14 - __ = 6	0.33	0.38	4.70	5.01	0.11
49	28 - 3 = __	0.35	0.17	5.20	4.26	5.82
50	29 - 10 = __	0.35	0.41	4.40	5.26	0.42
51	20 - __ = 12	0.38	0.43	5.50	5.38	0.68
52	27 - __ = 19	0.39	0.57	5.00	6.13	3.61
53	18 - __ = 8	0.40	0.50	6.10	5.53	0.57
54	17 - __ = 15	0.40	0.24	3.00	4.71	2.01
55	20 - 7 = __	0.42	0.33	6.10	4.93	1.70
56	28 - __ = 18	0.46	0.53	7.00	5.83	0.49
57	__ - 7 = 23	0.46	0.56	6.20	6.18	0.96
58	18 - __ = 11	0.47	0.31	6.20	4.81	1.65
59	17 - __ = 7	0.47	0.49	4.30	5.47	0.02

Table 15 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
60	__ - 9 = 8	0.53	0.38	6.00	5.02	1.48
61	__ - 4 = 7	0.53	0.17	6.90	3.97	13.78
62	__ - 6 = 19	0.54	0.40	6.20	5.42	2.07
63	__ - 9 = 16	0.54	0.53	6.40	5.83	0.02
64	__ - 10 = 19	0.58	0.47	5.90	5.58	1.10
65	__ - 2 = 17	0.60	0.17	5.00	4.20	19.78
66	__ - 4 = 8	0.60	0.18	5.50	4.03	18.35
67	__ - 10 = 9	0.60	0.38	6.10	4.96	3.20
68	__ - 7 = 5	0.60	0.26	8.20	4.44	8.74
69	__ - 10 = 6	0.60	0.35	5.00	4.77	4.18
70	__ - 7 = 21	0.62	0.34	7.00	5.11	8.55
71	__ - 7 = 17	0.62	0.43	6.00	5.50	3.58
72	__ - 9 = 19	0.65	0.56	7.30	6.01	1.03
73	18 - 5 = __	0.73	0.16	5.40	3.91	36.64
74	__ - 6 = 6	0.73	0.23	3.60	4.30	21.10
75	__ - 6 = 7	0.80	0.24	4.20	4.36	25.88
76	22 - 8 = __	0.85	0.39	6.80	5.19	22.97

χ^2 = 445.57 (76 items) χ^2(items < 10) = 136.30 (61 items) s^2 = 1.68

Table 16

Predicted and observed proportions of errors and success-latency in fifth-grade subtraction, level 1

Rank	Equations	Observed (1 - p_i)	Predicted (1 - p_i)	Observed Latency	Predicted Latency	X^2
1	18 - __ = 17	0.03	0.13	1.70	3.22	1.91
2	19 - __ = 19	0.05	0.11	2.20	3.14	0.70
3	17 - __ = 17	0.05	0.11	1.50	2.90	0.70
4	19 - __ = 18	0.05	0.13	2.00	3.34	1.10
5	20 - __ = 20	0.05	0.11	1.90	3.26	0.70
6	25 - 1 = __	0.05	0.10	2.20	3.41	0.24
7	22 - __ = 22	0.05	0.11	2.10	3.51	0.35
8	__ - 0 = 22	0.10	0.11	3.50	3.51	0.01
9	25 - 10 = __	0.10	0.38	3.70	5.18	3.24
10	__ - 1 = 17	0.15	0.10	4.00	2.57	0.70
11	18 - __ = 9	0.20	0.41	3.10	4.79	3.64
12	__ - 5 = 18	0.20	0.39	6.00	5.92	1.45
13	__ - 4 = 17	0.20	0.34	6.60	5.48	0.88
14	18 - 9 = __	0.25	0.41	3.20	4.79	2.11
15	19 - 2 = __	0.25	0.08	3.90	2.23	7.24
16	19 - __ = 9	0.25	0.46	3.50	5.11	3.47
17	18 - __ = 13	0.30	0.24	4.20	4.01	0.36
18	__ - 1 = 15	0.30	0.10	4.60	2.32	9.74
19	__ - 2 = 19	0.30	0.26	5.40	5.09	0.08
20	__ - 5 = 19	0.30	0.39	4.70	6.04	0.31
21	__ - 6 = 15	0.30	0.43	6.30	5.87	0.71
22	20 - __ = 13	0.35	0.48	4.30	5.95	1.36
23	19 - 3 = __	0.35	0.13	3.40	3.08	8.02
24	20 - 3 = __	0.35	0.23	4.40	4.51	1.53
25	__ - 3 = 22	0.40	0.13	6.00	3.81	6.08
26	__ - 10 = 12	0.40	0.46	5.60	5.47	0.13
27	__ - 3 = 17	0.45	0.30	6.10	5.16	2.19

Table 16 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
28	__ - 7 = 9	0.50	0.40	6.20	4.81	0.88
29	21 - 8 = __	0.50	0.44	7.40	5.61	0.12
30	21 - __ = 17	0.50	0.34	5.00	5.48	1.13
31	__ - 8 = 12	0.60	0.53	6.60	6.15	0.41
32	__ - 5 = 13	0.60	0.24	5.70	4.01	13.93
33	__ - 6 = 19	0.60	0.43	7.10	6.36	1.15
34	22 - __ = 15	0.60	0.48	5.60	6.19	0.57
35	24 - 7 = __	0.60	0.40	6.50	5.78	1.71
36	25 - 8 = __	0.60	0.44	6.80	6.10	0.97
37	__ - 9 = 14	0.70	0.58	7.70	6.71	0.62
38	__ - 8 = 16	0.70	0.53	6.80	6.63	1.18

$x^2 = 81.63$ (38 items) x^2(items < 10) = 67.70 (37 items) $s^2 = 1.64$

Table 17

Predicted and observed proportions of errors and success-latency in fifth-grade subtraction, level 2

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
1	44 - 22 = ___	0.03	0.22	3.00	4.55	4.30
2	48 - 22 = ___	0.13	0.21	4.70	4.67	0.70
3	47 - 20 = ___	0.15	0.20	4.90	4.68	0.34
4	58 - 35 = ___	0.15	0.26	5.30	4.72	1.24
5	46 - 32 = ___	0.18	0.26	5.00	4.42	1.66
6	48 - 36 = ___	0.20	0.28	4.40	4.40	1.47
7	48 - 37 = ___	0.20	0.29	3.60	4.38	1.66
8	55 - 32 = ___	0.20	0.25	4.40	4.69	0.25
9	58 - 33 = ___	0.20	0.25	5.80	4.70	0.29
10	48 - 20 = ___	0.20	0.20	4.40	4.71	0.00
11	37 - 26 = ___	0.24	0.25	4.10	4.26	0.00
12	37 - 25 = ___	0.24	0.24	4.00	4.28	0.00
13	47 - 32 = ___	0.24	0.26	5.30	4.45	0.07
14	37 - 26 = ___	0.24	0.25	3.30	4.26	0.00
15	44 - 20 = ___	0.25	0.21	5.40	4.59	0.23
16	53 - 32 = ___	0.25	0.25	4.60	4.63	0.00
17	43 - 21 = ___	0.25	0.21	4.50	4.54	0.14
18	57 - 33 = ___	0.25	0.25	5.30	4.73	0.00
19	56 - 30 = ___	0.25	0.24	5.20	4.75	0.02
20	48 - 33 = ___	0.27	0.26	5.90	4.46	0.00
21	41 - 30 = ___	0.27	0.26	4.30	4.31	0.01
22	38 - 27 = ___	0.29	0.25	4.30	4.28	0.38
23	46 - 23 = ___	0.31	0.22	4.10	4.59	0.86
24	48 - 31 = ___	0.33	0.25	5.20	4.50	1.48
25	58 - 31 = ___	0.38	0.24	4.40	4.80	1.64
26	56 - 32 = ___	0.38	0.25	3.70	4.72	1.42
27	51 - 25 = ___	0.38	0.68	4.90	6.67	6.80
28	36 - 21 = ___	0.40	0.22	5.40	4.33	8.22

Table 17 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
29	50 - 24 = ___	0.40	0.68	6.30	6.66	6.90
30	50 - 24 = ___	0.44	0.68	4.80	6.66	4.10
31	55 - 34 = ___	0.44	0.26	4.00	4.65	2.68
32	40 - 18 = ___	0.45	0.66	6.50	6.48	3.80
33	40 - 18 = ___	0.50	0.66	6.40	6.48	2.18
34	44 - 21 = ___	0.56	0.21	3.10	4.57	11.92
35	57 - 28 = ___	0.56	0.69	5.40	6.80	1.13
36	53 - 24 = ___	0.56	0.67	7.30	6.75	0.82
37	40 - 19 = ___	0.56	0.66	6.10	6.46	0.72
38	40 - 17 = ___	0.56	0.65	6.60	6.50	0.54
39	31 - 16 = ___	0.60	0.66	6.80	6.25	0.76
40	41 - 15 = ___	0.60	0.64	7.30	6.57	0.11
41	41 - 15 = ___	0.63	0.64	7.10	6.57	0.01
42	42 - 13 = ___	0.63	0.62	6.90	6.64	0.00
43	43 - 14 = ___	0.63	0.63	6.60	6.65	0.00
44	54 - 26 = ___	0.63	0.68	7.30	6.75	0.22
45	52 - 24 = ___	0.65	0.67	8.30	6.72	0.04
46	31 - 18 = ___	0.69	0.67	6.20	6.21	0.05
47	50 - 24 = ___	0.70	0.68	6.80	6.66	0.06
48	32 - 15 = ___	0.71	0.65	5.50	6.30	0.66
49	43 - 14 = ___	0.75	0.63	7.70	6.65	1.32
50	42 - 15 = ___	0.75	0.63	6.70	6.60	1.16
51	41 - 27 = ___	0.76	0.71	7.00	6.34	0.49
52	43 - 24 = ___	0.78	0.69	6.00	6.46	1.71
53	44 - 27 = ___	0.78	0.70	5.70	6.43	1.20
54	43 - 26 = ___	0.82	0.70	6.70	6.42	3.21
55	53 - 28 = ___	0.85	0.69	---	---	2.31
56	55 - 29 = ___	0.85	0.70	7.50	6.72	2.25
57	51 - 23 = ___	0.95	0.67	---	---	7.21

$X^2 = 90.67$ (57 items) $X^2(\text{items} < 10) = 78.76$ (56 items) $S^2 = 0.64$

Table 18

Predicted and observed proportions of errors and success-latency in fifth-grade subtraction, level 3

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
1	49 - 20 = __	0.03	0.10	3.40	3.12	0.80
2	47 - 24 = __	0.05	0.16	2.70	3.41	0.88
3	46 - 21 = __	0.07	0.13	3.90	3.25	0.55
4	57 - 32 = __	0.07	0.18	5.20	3.71	1.23
5	49 - 25 = __	0.07	0.15	4.10	3.44	0.87
6	46 - 24 = __	0.07	0.17	3.70	3.43	1.07
7	57 - 35 = __	0.10	0.22	3.00	3.90	0.83
8	47 - 21 = __	0.10	0.12	3.00	3.23	0.06
9	45 - 20 = __	0.10	0.13	2.70	3.20	0.07
10	49 - 27 = __	0.10	0.18	2.20	3.56	0.43
11	67 - 30 = __	0.10	0.09	2.90	3.39	0.01
12	51 - 20 = __	0.10	0.09	2.90	3.08	0.00
13	58 - 27 = __	0.10	0.09	3.20	3.30	0.01
14	59 - 21 = __	0.10	0.06	3.20	3.00	0.10
15	56 - 25 = __	0.10	0.11	3.10	3.30	0.00
16	59 - 21 = __	0.10	0.06	3.90	3.00	0.10
17	47 - 23 = __	0.13	0.15	4.40	3.35	0.02
18	59 - 37 = __	0.20	0.23	3.50	3.99	0.05
19	54 - 31 = __	0.20	0.19	2.80	3.71	0.01
20	47 - 22 = __	0.20	0.14	3.80	3.29	0.36
21	53 - 30 = __	0.20	0.18	1.90	3.67	0.02
22	66 - 33 = __	0.20	0.12	2.50	3.60	0.31
23	62 - 27 = __	0.20	0.43	6.70	6.53	1.08
24	60 - 23 = __	0.20	0.33	8.10	6.26	0.38
25	69 - 38 = __	0.20	0.15	2.30	3.85	0.09
26	50 - 12 = __	0.20	0.25	7.00	5.78	0.06
27	50 - 12 = __	0.20	0.25	4.80	5.78	0.06

Table 18 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	X^2
28	62 - 26 = ___	0.20	0.37	7.00	6.41	0.59
29	56 - 17 = ___	0.20	0.27	5.00	5.97	0.11
30	44 - 33 = ___	0.24	0.34	3.20	4.03	3.15
31	50 - 21 = ___	0.27	0.43	5.50	6.34	1.61
32	34 - 22 = ___	0.27	0.26	4.20	3.55	0.06
33	30 - 15 = ___	0.28	0.59	3.00	6.36	28.22
34	47 - 30 = ___	0.30	0.25	4.60	3.79	0.99
35	45 - 21 = ___	0.30	0.14	2.50	3.27	2.18
36	50 - 24 = ___	0.30	0.50	4.50	6.52	1.55
37	57 - 36 = ___	0.30	0.24	3.80	3.96	0.23
38	51 - 26 = ___	0.30	0.53	7.80	6.63	2.08
39	44 - 16 = ___	0.30	0.41	5.50	6.14	0.46
40	39 - 26 = ___	0.31	0.27	5.10	3.70	0.66
41	49 - 30 = ___	0.35	0.22	3.60	3.75	6.75
42	49 - 35 = ___	0.37	0.31	4.80	4.06	0.94
43	49 - 34 = ___	0.37	0.29	4.60	4.00	1.81
44	49 - 30 = ___	0.39	0.22	4.30	3.75	11.89
45	50 - 29 = ___	0.40	0.51	6.50	6.84	2.77
46	47 - 18 = ___	0.40	0.41	7.30	6.21	0.00
47	59 - 30 = ___	0.40	0.14	4.30	3.55	5.92
48	42 - 16 = ___	0.40	0.44	5.40	6.18	0.05
49	56 - 19 = ___	0.40	0.30	7.20	6.09	0.22
50	62 - 29 = ___	0.40	0.43	6.90	6.60	0.02
51	55 - 16 = ___	0.40	0.26	4.00	5.93	0.51
52	50 - 12 = ___	0.40	0.25	3.60	5.78	0.62
53	65 - 27 = ___	0.40	0.34	7.60	6.41	0.07
54	49 - 30 = ___	0.47	0.22	4.60	3.75	23.85
55	46 - 17 = ___	0.47	0.40	6.00	6.17	0.29
56	41 - 16 = ___	0.47	0.45	7.40	6.20	0.02
57	52 - 26 = ___	0.50	0.51	5.80	6.61	0.01

Table 18 (contimued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	x^2
58	45 - 19 = __	0.50	0.46	6.50	6.31	0.07
59	40 - 21 = __	0.52	0.58	5.20	6.54	0.97
60	41 - 13 = __	0.53	0.38	8.50	6.02	1.42
61	43 - 19 = __	0.53	0.49	6.50	6.35	0.12
62	37 - 19 = __	0.58	0.58	6.60	6.47	0.00
63	46 - 27 = __	0.58	0.62	5.50	6.79	0.65
64	36 - 17 = __	0.59	0.56	6.50	6.39	0.25
65	53 - 29 = __	0.60	0.57	9.50	6.78	0.07
66	54 - 26 = __	0.60	0.48	6.50	6.57	0.83
67	43 - 17 = __	0.60	0.44	7.50	6.23	1.50
68	44 - 18 = __	0.60	0.45	6.30	6.27	1.35
69	51 - 25 = __	0.60	0.50	8.10	6.57	0.36
70	46 - 28 = __	0.65	0.64	7.20	6.85	0.00
71	32 - 13 = __	0.66	0.52	6.00	6.20	5.98
72	37 - 19 = __	0.69	0.58	7.20	6.47	3.64
73	45 - 27 = __	0.72	0.64	6.90	6.81	2.00
74	51 - 28 = __	0.73	0.57	6.90	6.75	1.57
75	55 - 29 = __	0.73	0.54	4.80	6.74	2.34
76	31 - 17 = __	0.78	0.62	6.50	6.46	7.15

$x^2 = 137.34$ (76 items) x^2(items < 10) $= 73.39$ (73 items) $s^2 = 1.25$

Table 19

Predicted and observed proportions of errors and success-latency in fifth-grade subtraction, level 4

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
1	29 - 5 = HK	0.04	0.18	3.60	4.56	1.52
2	29 - 24 = M	0.04	0.29	1.20	4.67	3.61
3	49 - H = 43	0.06	0.21	3.20	5.17	1.20
4	36 - 5 = MK	0.07	0.16	4.40	4.53	0.43
5	43 - 21 = R2	0.07	0.16	3.10	3.87	0.40
6	42 - 5 = R7	0.08	0.22	5.30	5.20	1.29
7	25 = 29 - __	0.11	0.25	2.60	5.26	0.92
8	38 - 12 = M6	0.11	0.13	3.50	3.85	0.03
9	24 = 28 - __	0.14	0.25	2.50	5.27	0.44
10	37 - M = 31	0.14	0.17	3.50	4.53	0.03
11	41 = 49 - __	0.17	0.22	4.40	5.18	0.18
12	37 = 42 - __	0.22	0.41	5.40	6.61	1.33
13	__ - 5 = 23	0.22	0.26	7.70	5.27	0.06
14	37 - 5 = HK	0.22	0.11	4.20	3.81	1.22
15	32 - __ = 25	0.22	0.47	6.10	6.67	2.16
16	37 - K = 32	0.25	0.23	3.70	5.23	0.02
17	40 - __ = 26 - 25	0.25	0.71	5.90	8.71	12.60
18	43 - 5 = __	0.29	0.30	7.70	5.90	0.01
19	__ - 6 = 24	0.33	0.47	5.40	6.68	0.64
20	33 = __ - 6	0.33	0.23	5.60	5.22	0.51
21	54 - 17 = F7	0.33	0.26	6.30	5.20	0.27
22	43 - 8 = __	0.33	0.33	6.90	5.91	0.00
23	__ - 5 = 29	0.42	0.24	7.00	5.24	2.04
24	__ - 9 = 25	0.42	0.48	5.60	6.67	0.16
25	32 - __ = 27	0.42	0.45	5.20	6.66	0.05
26	27 = __ -	0.42	0.45	6.20	6.66	0.04
27	__ - 7 = 19	0.43	0.49	8.20	6.71	0.10
28	27 = __ - 4	0.43	0.45	7.80	6.66	0.01

Table 19 (continued)

Rank	Equations	Observed $(1 - p_i)$	Predicted $(1 - p_i)$	Observed Latency	Predicted Latency	χ^2
29	28 - KR = 7	0.43	0.37	4.30	5.36	0.09
30	26 = __ - 7	0.44	0.46	8.00	6.67	0.01
31	39 - RN = 7	0.44	0.42	5.80	5.36	0.02
32	__ = 35 - 8	0.44	0.46	8.30	6.66	0.01
33	42 - 40 = 38 - __	0.44	0.69	4.10	6.80	2.65
34	35 = 41 - __	0.50	0.42	4.50	6.62	0.29
35	__ - 25 = 37 - 36	0.50	0.44	5.90	6.67	0.18
36	38 - 9 = __	0.50	0.35	4.10	5.95	1.19
37	28 = __ - 5	0.50	0.24	6.20	5.25	4.29
38	__ = 26 - 9	0.56	0.51	8.90	6.73	0.06
39	26 - __ = 30 - 6	0.56	0.67	6.90	8.09	0.53
40	__ - 26 = 32 - 30	0.57	0.64	8.60	6.80	0.14
41	27 = __ - 8	0.57	0.46	8.80	6.66	0.33
42	__ = 34 - 8	0.57	0.47	9.20	6.67	0.31
43	17 - 13 = 26 - __	0.57	0.57	4.60	6.79	0.00
44	27 - 24 = 33 - __	0.57	0.63	6.50	6.79	0.09
45	__ = 32 - 5	0.57	0.45	7.50	6.66	0.42
46	__ - 24 = 6	0.57	0.61	8.00	6.78	0.06
47	34 - 31 = __ - 26	0.58	0.54	8.50	7.35	0.09
48	__ = 43 - 7	0.67	0.42	8.50	6.61	2.90
49	45 - FG = 3	0.67	0.48	6.50	5.38	1.66
50	42 - 15 = K7	0.71	0.28	6.00	5.25	6.48
51	__ = 32 - 6	0.75	0.46	9.00	6.67	4.12
52	__ - 27 = 36 - 32	0.78	0.87	---	---	0.73
53	53 - 17 = M6	0.83	0.36	7.10	5.91	11.78
54	23 - 8 = MH	0.86	0.40	9.60	6.02	6.19
55	27 - __ = 32 - 6	0.86	0.88	8.40	10.19	0.05
56	25 - 21 = __ - 33	0.89	0.80	8.80	8.20	0.48
57	24 - __ = 16	0.93	0.51	---	---	5.01

$\chi^2 = 81.43$ (55 items) χ^2(items < 10) = 57.05 (55 items) $s^2 = 2.93$

obtained for addition. In the case of Tables 17 and 19 the X^2 values, including the largest individual item contributions are just significant at the .01 level and not significant at all in the case of Table 13.

Tables 13-19 also show the predicted and observed success-latencies for the subtraction data. Again the statistic S^2 is given at the bottom of each table. Applying the same interpretation as before to this statistic, we may look at the value of S for each table as an estimate of the standard deviation of the approximately normal distribution of errors. For these seven tables we find the values of S to be 1.96, 1.34, 1.30, 1.28, 0.80, 1.12 and 1.71 respectively, and it is reasonable to say that errors of prediction greater than about 1.50 seconds should not occur very often. The observed success-latency values have a range from slightly more than 5 seconds (Table 17) to more than 8 seconds (Table 19), and consequently, a model with errors that have an approximately normal distribution with a standard deviation of about 1.5 seconds yields meaningful and useful predictions. These results are very comparable to those found for addition. The same is true of the measure of average percentage error, which is 25.7%, 23.0%, 24.6%, 26.3%, 10.9%, 17.1%, and 23.0% for Tables 13-19 respectively.

Without making an exact statistical comparison it still seems clear that the approximate measures of fit we have reported for the success-latencies in subtraction reflect a better fit to the data than do the X^2 measures for predicted response proportions. The predictions of response proportions still leave a lot to be desired. The predictions of success-latencies seem to reflect more regularly the observed rankings of latencies, even though this apparent difference in favor of latency predictions is not well reflected in the multiple correlation coefficients of Table 12.

Inspection of the tables for subtraction confirms the intuition that subtraction problems of the form $\_\_ - n = p$ are not relatively as difficult as the same form is in the case of addition. No doubt the reason for this is that a single simple transformation converts such subtraction problems into the easiest sort of addition problem, $p + n = \_\_$.

It should be noted that Table 19 includes problems using letter variables as well as blanks, and it is interesting to note that problems using letter variables are the six easiest problems in the table in terms of response errors, although the same six problems do not have the shortest latencies.[4] The format of these problems with letter variables was of the following sort:

$$\begin{array}{r} 45 \\ -FG \\ \hline 3 \end{array} \quad FG = \_\_\_ \qquad \begin{array}{r} 29 \\ -\ 5 \\ \hline HK \end{array} \quad HK = \_\_\_ \qquad \begin{array}{r} 37 \\ -\ K \\ \hline 32 \end{array} \quad K = \_\_\_ \qquad \begin{array}{r} 29 \\ -24 \\ \hline M \end{array} \quad M = \_\_\_ .$$

The ease of handling algebraic notation is also confirmed by some other unpublished experiments conducted in the Institute several years ago with first- and second-grade children.

Multiplication--grade four. The two sets of problems considered each contained twenty exercises, as in the case of addition and subtraction. The two sets concentrated on a review of multiples of 4 and 5, with the second factor ranging from 0 to 12. The problems occurred in the three forms, $m \times n = \_\_$, $m \times \_\_ = p$, and $\_\_ \times n = p$. Unlike the addition and subtraction analyses covered in the previous pages NSTEPS was not considered as a variable because in all problems only one operation was involved. To see if transformations as described in the theory section defined a significant variable, we treated each of the three equational forms as an independent variable which took on the value 1 if the problem was in the given form and 0 if it were not. The other two independent variables used were the larger factor (LARGER) and smaller factor (SMALLER) that yielded the product. In the case of squares (4×4 and 5×5) the values of the two factors were equal. Table 20 presents the regression coefficients for the five variables considered with proportion of errors and success-latency as dependent variables.

Insert Table 20 about here

Again we found that the linear-regression model does well at predicting errors and success-latency from a small number of variables. The only equation-form variable that significantly affected the regression line was the canonical form $a \times b = \_\_\_$ and the negative coefficients of this variable indicate that problems of this form are easier than problems of the form $\_\_\_ \times b = c$ or $a \times \_\_\_ = c$, a finding well in keeping with intuition.

When all forty problems were considered the overall $\chi^2 = 113.29$, but the deletion of three problems each with a χ^2 component greater than 10 dropped the total for the remaining 37 problems to $\chi^2 = 41.43$. The three problems dropped from the analysis were $4 \times 10 = \_\_\_$, for which the predicted error proportion was much higher than observed, and $4 \times \_\_\_ = 4$ and $\_\_\_ \times 4 = 48$ for which the predicted error proportions were much

Table 20

Regression coefficients for fourth-grade multiplication

Dependent	Problems	Subjects	Constant	Larger	Smaller	a × b = __	a × __ = c	__ × a = c	R	R^2
Errors	40	47	-1.46	0.10	0.01	-0.29	----	----	0.70	0.50
Success-latency	40	47	1.02	0.18	0.33	-0.38	----	----	0.78	0.62

lower than observed. It is not difficult to analyze why these three problems probably deviated greatly from the predicted values. In general $4 \times 10 = \_\_$ allows use of the simple algorithm $a \times 10 = a0$. We would expect the same low error finding for $5 \times 10 = \_\_$, but this problem did not occur in the two sets. The problem $4 \times \_\_ = 4$ turned out to be the first multiplication problem presented, and as mentioned in the discussion of addition, there is evidence of a warm-up effect which affects response to the first problem of the day. The problem $\_\_ \times 4 = 48$ is the only problem of the set for which the initial factor is both 12 and also the answer to be found. The s^2 for comparing observed and predicted latency was quite low. The obtained value, $s^2 = .62$, indicates that most prediction errors were definitely less than 1.0 second.

Multiplication tables--grades three, four, five and six. Toward the end of the school year we decided to run the 100 one-digit multiplication problems of the form $a \times b = \_\_$ to see how well a structural model would predict response behavior. Previous investigations of performance on these basic multiplication facts are not as numerous as we had expected, and the kind of regression model applied here has not been previously used, as far as we know. The first point to note is that for all four grades, the response performance was extremely good. The error rate was 8.0 per cent for third-grade children and 3.2 per cent for sixth-grade children, with the fourth and fifth grades falling between these two bounds. Consequently our analysis in this case is restricted entirely to success-latencies. Because the form of the equations was constant in the 100 problems, we have restricted our regression to the two factors, SMALLER and LARGER, already used in analyzing fourth-grade multiplication. The regression coefficients, multiple correlation and statistic S^2 for each grade are shown in Table 21. There are several observations to be made about this

----

Insert Table 21 about here

----

table. In the first place, for all four grades the multiple correlation R is extremely high, indicating that the two variables are giving a good account of the data. This inference is supported by the small values of S^2, which are the lowest values reported for any of the sets of data analyzed in this paper. It is also apparent from the values of the regression coefficients that the magnitude of the smaller factor is more important than that of the larger factor. Thus, for example, on the average it takes longer to say what 1×9 is than to say what 4×5 is. Finally, with analysis for four grades before us, it is natural to ask whether we can find evidence of development from one grade to another. Development is most evident in the monotonically decreasing values of the constant, which reflect an increase in speed of response with age. In the regression model for latencies, the constant enters in a direct additive way. The decrease from 1.71 seconds in the third grade to 1.33 seconds in the sixth grade is not surprising. What is surprising is that the coefficients of the two factors do not show a corresponding monotonicity with age. This lack of monotonicity complicates considerably the task of constructing a model of developmental processes and their effects on arithmetic performance.

Figures 5 and 6 show the predicted and observed success-latency curves for the third and sixth grades respectively. The 100 problems

----

Insert Figures 5 and 6 about here

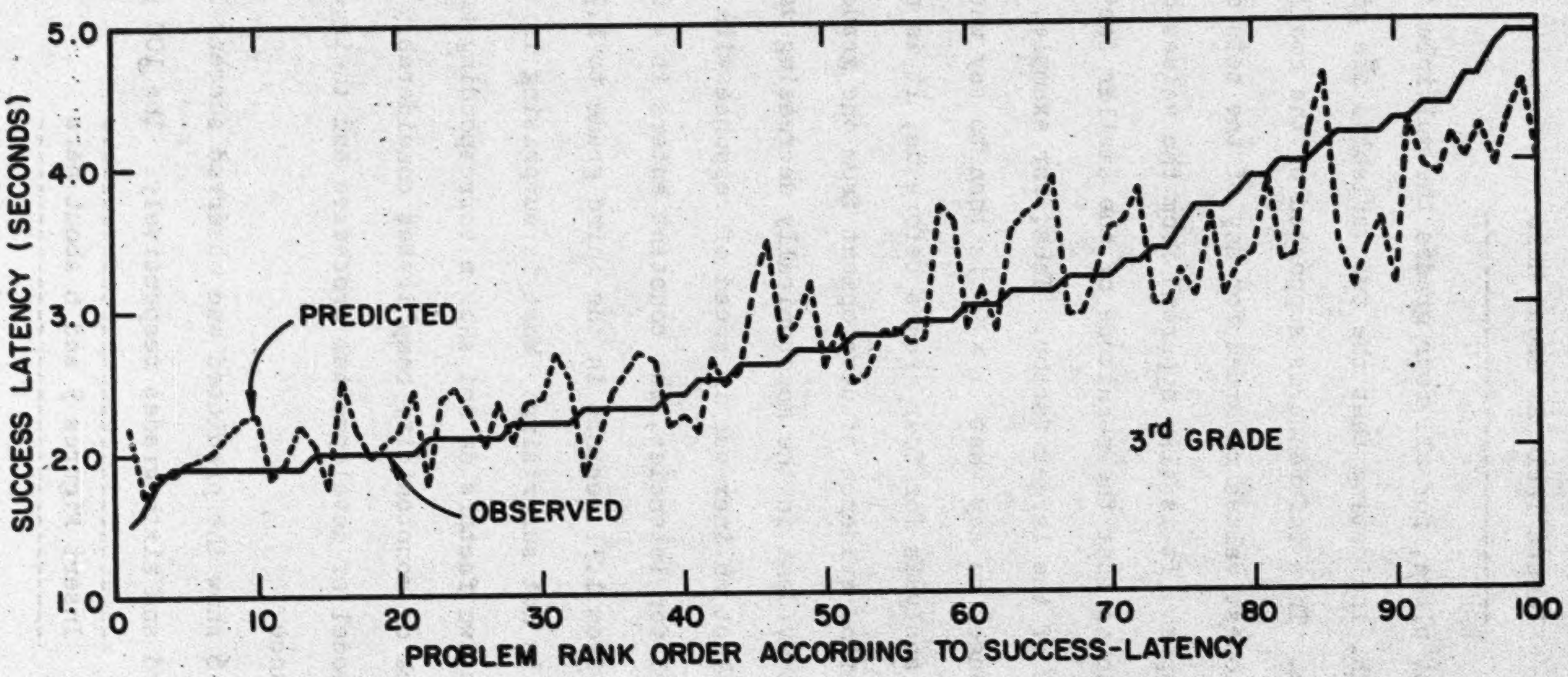

Figure 5. Predicted and observed success-latencies of third-grade students for the multiplication tables.

Table 21

Linear-regression coefficients for the multiplication tables

Grade	Subjects	Constant	Larger Factor	Smaller Factor	R	R^2	S^2
3	24	1.71	0.06	0.30	0.86	0.74	0.22
4	56	1.52	0.07	0.28	0.85	0.73	0.22
5	20	1.38	0.09	0.29	0.78	0.61	0.42
6	32	1.33	0.06	0.19	0.82	0.68	0.14

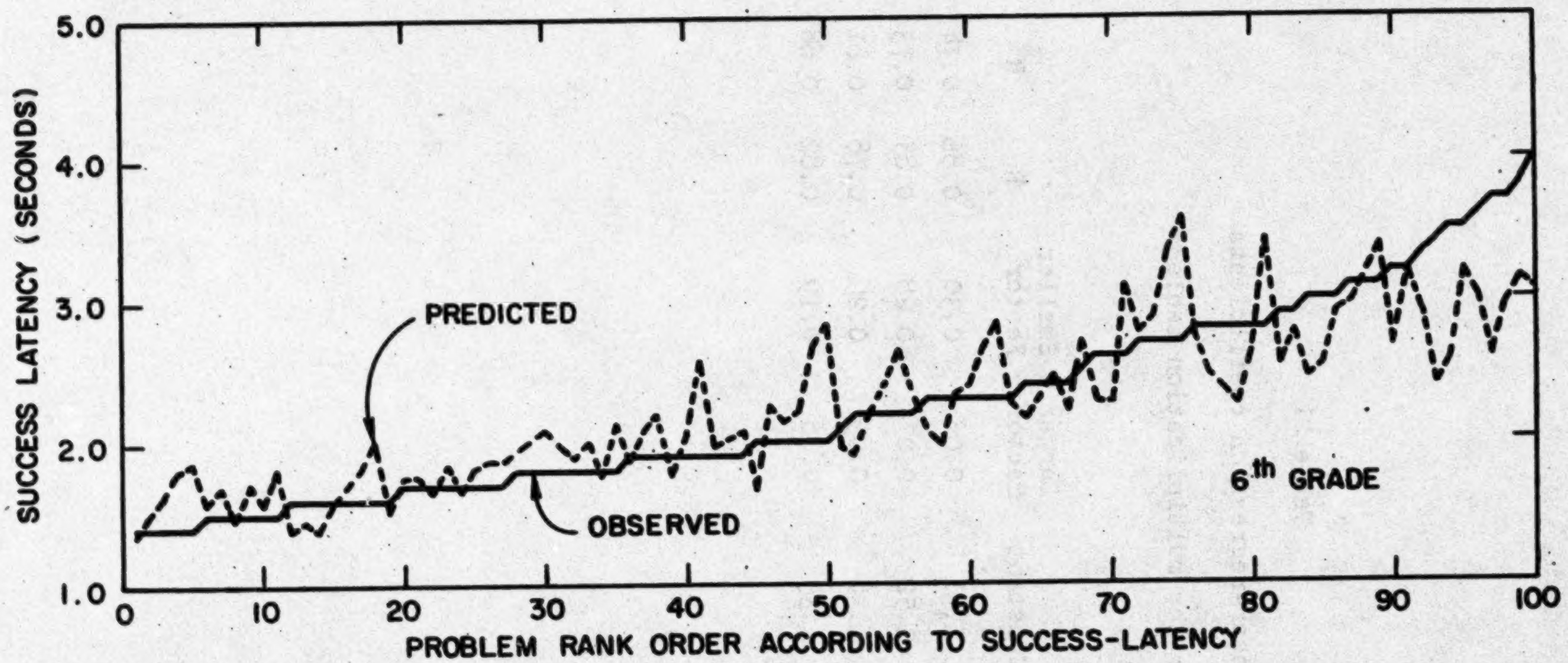

Figure 6. Predicted and observed success-latencies of sixth-grade students for the multiplication tables.

Table 22

Linear-regression coefficients for transformation, operation and memory steps in fourth-grade addition

	Constant	Transformation	Operation	Memory	R	R^2
Errors	-1.29	0.20	0.00	0.39	0.73	0.53
Latencies	2.94	0.58	0.00	0.75	0.69	0.48

are rank ordered according to success-latency on the abscissa, and thus the observed data define a relatively smooth monotonically increasing function. The predicted curve is determined for each grade level by the three estimated coefficients given in Table 21 and the two given factors of each multiplication problem. Considering the wide range of latencies found in each figure, running from 1.5 to 4.9 seconds in the third grade, and from 1.4 to 4.0 seconds in the sixth grade, we feel that the predicted curves are fitting the observed data quite well. For those readers accustomed to looking at smooth predicted learning curves that are essentially exponential in form, we emphasize that the predictive task is different and rather more difficult, as we move not from like trial to like trial, but from a problem-item with a particular structure to another problem-item with a distinct structure.

Analysis of the factors in NSTEPS. As we promised in the theoretical discussion of the second section, we now present a preliminary analysis of breaking up the single variable NSTEPS into its three components of *transformation*, *operation* and *memory*. There is one slight difference in the analysis presented here from the definition given in the second section. Transformation steps always were either 0 or 1, never 2. With this exception, the analysis was entirely based on the earlier definitions. The data used in the first analysis are those already reported in Tables 5 and 6, but without the first item of each set of problems deleted. Thus this first analysis is in terms of 80 fourth-grade addition problems. The results are shown in Table 22. In the case of both errors and

----

Insert Table 22 about here

----

success-latencies it is important to observe that memory is the most important variable, while operation plays no role. Moreover, in both cases we get nearly as good a fit simply by using memory as the single variable. In the case of errors the difference in the multiple correlation R occurs only in the third decimal, .726 rather than .731, and in the case of latency .677 rather than .688. The X^2 and S^2 values that come from using the coefficients of Table 22 are high but are not out of line with those reported earlier. In particular $X^2 = 417.7$, and if we delete the 12 extreme items having individual X^2's greater than 10, $X^2 = 170.3$, for the remaining 68 items. The statistic $S^2 = 1.42$, which yields an estimate of 1.19 for the standard deviation of the errors in prediction. What is particularly worth noting in a comparison of Tables 5 and 22 is that the correlation for fourth-grade addition (block 1, level 3) is lower than the correlation for the combined data of Table 22. (For the problems of this block, see Table 5.)

A second, somewhat different analysis was performed on a set of 19 problems that, together with the initial problem omitted in the analysis, formed one day's exercises on fourth-grade addition, block 3, level 4. These 19 problems are among the 76 already analyzed in Table 10. The departures from the earlier definitions of the components of NSTEPS were these. First, because the problems were all of the form ab + cd = __ + ef or ab + cd = ef + __, the number of transformations was the same for all problems and therefore was omitted as a variable. Second, the operations of addition and subtraction of single digits were treated as separate variables. Third, the number of digits in memory was expanded to include all digits used in obtaining a solution, including those presented in the problem, those that occurred as partial solutions, and those that were present in the response. The three variables considered were, therefore, number of addition operations (OP1), number of subtractions operations (OP2) and number of digits processed (MEMORY).

Table 23 presents the regression coefficients for the three variables found, with proportion of errors and success-latency as dependent variables.

Insert Table 23 about here

The very high correlations for both errors and latencies warrant a closer look at the results.

For the data entering this analysis the mean number of addition operations was 2.4, the mean number of subtraction operations was 1.8 and the mean number of digits processed was 8.7. It would appear that the number of addition operations has a much smaller effect on errors than the number of subtraction operations. Neither of these two variables has a significant effect on success-latency. Figure 7 presents the observed and predicted proportion of errors as a function of ranked difficulty. With the exception of problems 6 and 8, the observed and predicted curves are quite similar. Figure 8 presents observed and predicted success-latencies as a function of observed latency rank. Once more we find the general shapes of the observed and predicted curves quite similar. Figure 9 is a scatter plot of observed versus predicted errors. Figure 10 is a similar scatter plot of observed versus predicted success-latency. If

----

Insert Figures 7 and 8 about here

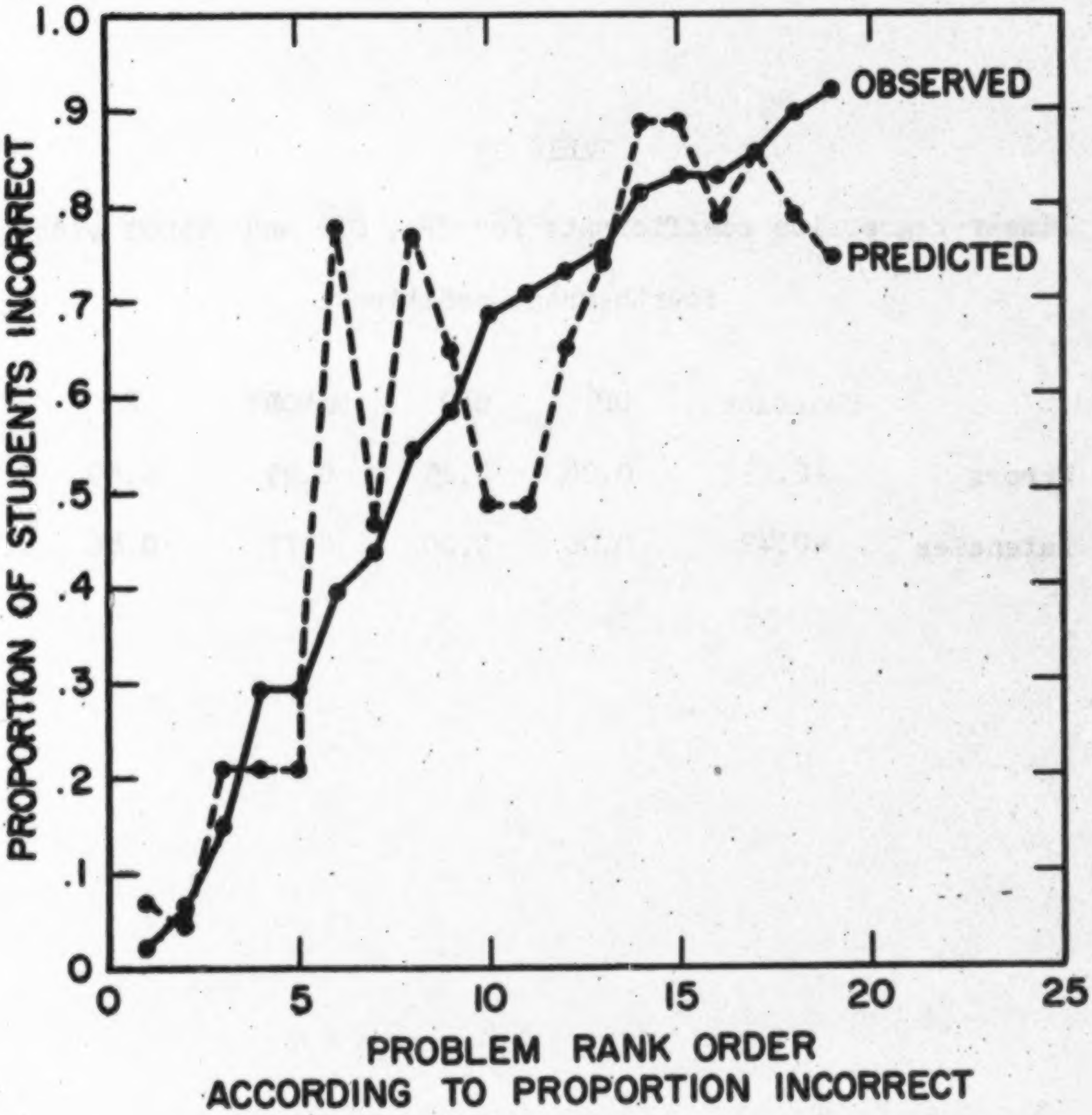

Figure 7. Predicted and observed proportions of errors on fourth-grade addition analyzed in terms of three process variables.

Table 23

Linear-regression coefficients for OP1, OP2 and MEMORY steps in fourth-grade addition

	Constant	OP1	OP2	MEMORY	R	R^2
Errors	-2.65	0.06	0.25	0.25	0.89	0.79
Latencies	-0.42	0.00	0.00	0.77	0.86	0.73

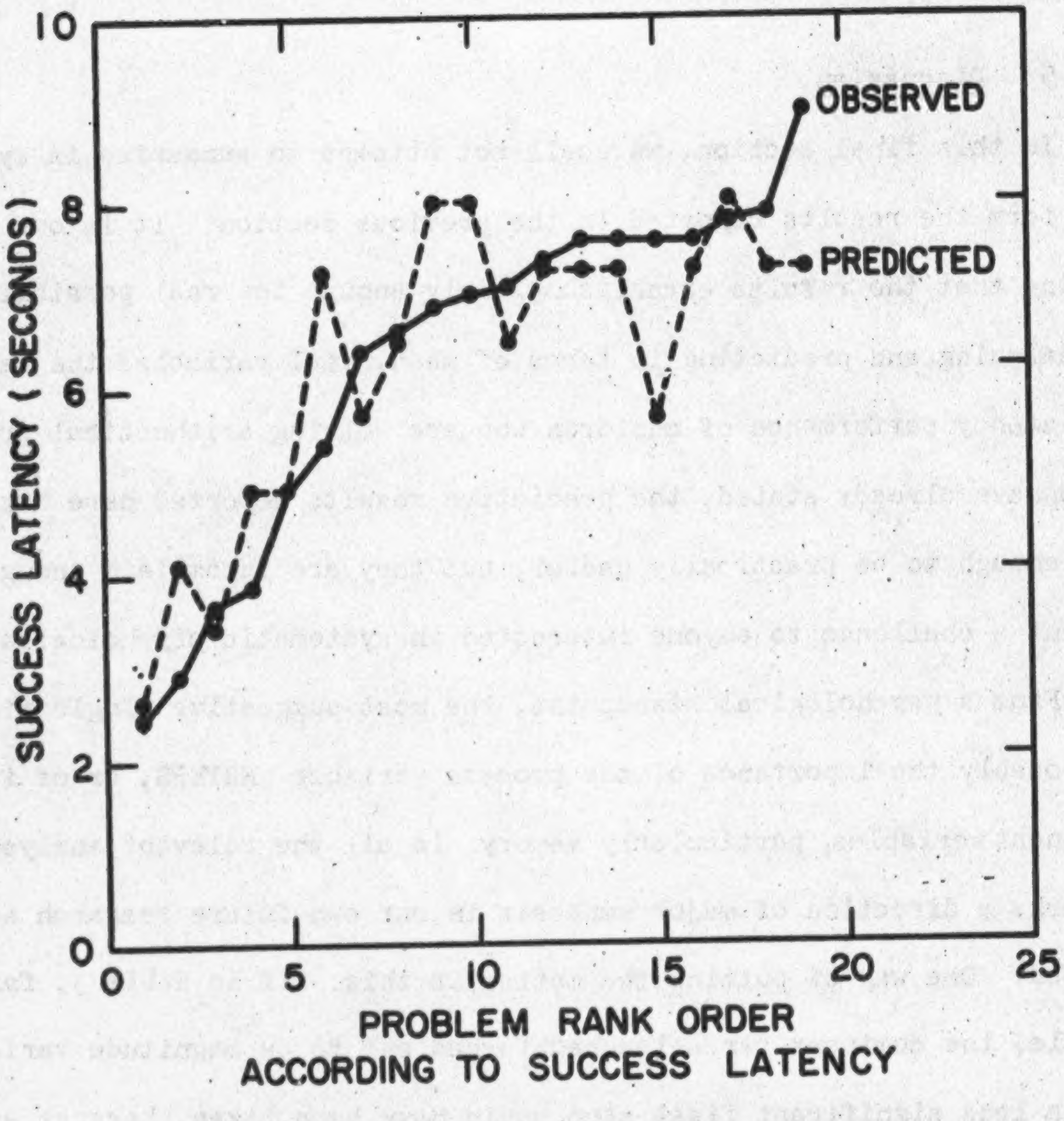

Figure 8. Predicted and observed success-latencies for fourth-grade addition analyzed in terms of three process variables.

Insert Figures 9 and 10 about here

---

all the points in the two plots fell on the 45° straight line the predictions would be perfect. The deviations of the points from this line are a measure of the goodness of fit of the model.

5. Discussion.

In this final section, we shall not attempt to summarize in systematic form the results reported in the previous section. It is our own feeling that the results establish clearly enough the real possibility of analyzing and predicting in terms of meaningful variables the response and latency performance of children who are solving arithmetical problems. As we have already stated, the predictive results reported here have been good enough to be practically useful, but they are incomplete enough to present a challenge to anyone interested in systematic psychological theory.

From a psychological standpoint, the most suggestive single finding is probably the importance of the process variable NSTEPS, or of its component variables, particularly memory, in all the relevant analyses. It marks a direction of major emphasis in our own future research as now planned. One way of putting the matter is this. If in Table 3, for example, the dominant variables had turned out to be magnitude variables, then a less significant first step would have been taken, because anyone would immediately ask what characteristics of the processing done internally by the students made these magnitude variables so significant. In postulating process variables and being able to establish their direct importance, we have already been able to move past this first step. Now our

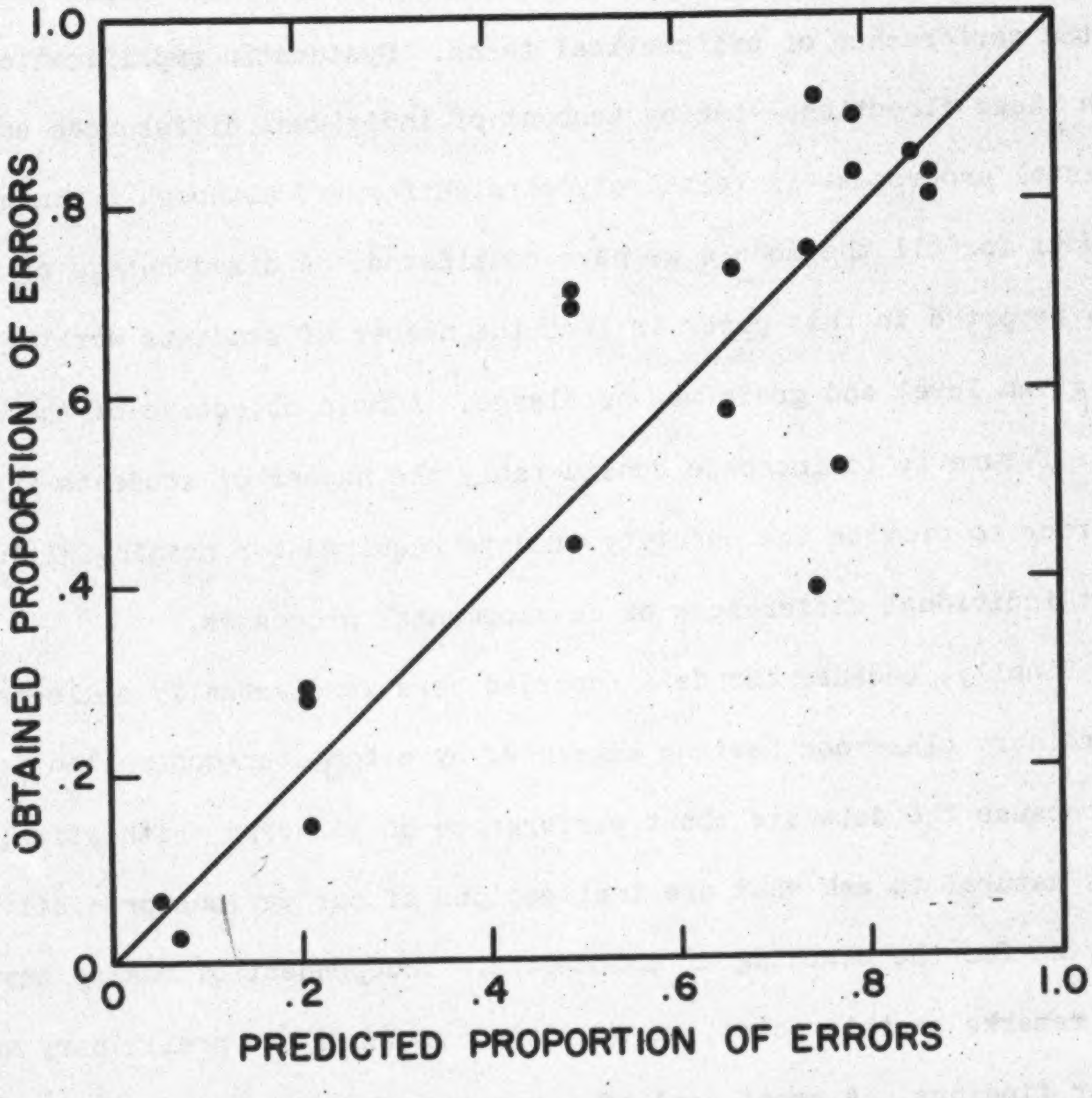

Figure 9. Scatter plot of observed versus predicted errors on fourth-grade addition.

central problem is to acquire a better understanding of these variables and to use this understanding to develop better predictive models.

All the analyses reported in this paper have been concerned with mean data averaged over individual student performance. Moreover, when dealing with data from different age groups, no attempt has been made to estimate parameters that would reflect the course of developmental change in the performance of arithmetical tasks. Systematic amplification in both these directions--taking account of individual differences and developmental processes--is relatively straightforward although technically arduous for all the models we have considered. A disadvantage of the data reported in this paper is that the number of students working at any given level and grade was not large. A main objective of the immediate future is to increase considerably the number of students involved in order to provide the quantity of data required for meaningful inferences about individual differences or developmental processes.

Finally, because the data reported here were actually collected in an ordinary classroom setting augmented by a computer-controlled terminal, and because the data are about performance on standard arithmetical problems, it is natural to ask what are implications of our various predictive analyses for the teaching of arithmetic. Independent of making any positive remarks on this point, we want to underscore the preliminary value of our findings. A great deal of more refined analysis with data from larger numbers of students is needed to support any definitive pedagogical recommendations. Keeping in mind this explicit reservation, we do feel that the results that are most intriguing from a pedagogical standpoint are the ones reported at the end of the last section on the ability of

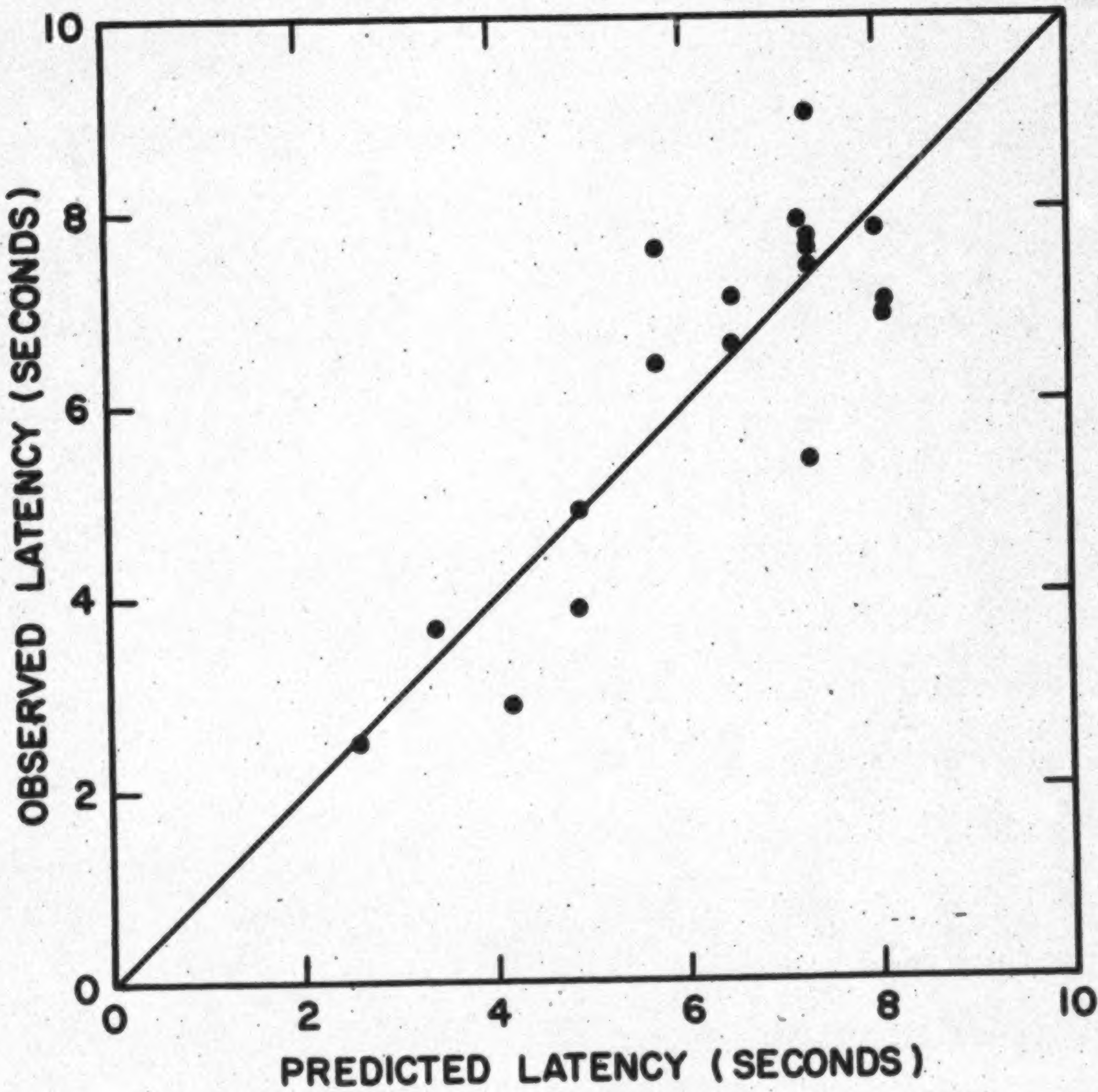

Figure 10. Scatter plot of observed versus predicted success-latencies for fourth-grade addition.

the memory variable alone to offer a fairly adequate account of the observed data. From the way this variable was defined in the theoretical section, it should be evident that we can identify some specific points to emphasize in teaching multi-digit addition and subtraction. However, we leave for another time and place the taking of this explicit pedagogical step.

References

Atkinson, R. C. and Estes, W. K. Stimulus sampling theory. In R. R. Bush, E. Galanter and R. D. Luce (Eds.), Handbook of Mathematical Psychology, Vol. II. New York: Wiley, 1963. Pp. 121-268.

Batson, W. H. and Combellick, O. E. Relative difficulty of number combinations in addition and multiplication. J. educ. Psychol., 1925, 26, 467-481.

Billington, Lillian E. A study of the time interval between stimulus and response in arithmetic fundamentals. Unpublished doctoral dissertation, Stanford Univer., 1947.

Bruner, J. S., Goodnow, Jacqueline J., and Austin, G. A. A Study of Thinking. New York: Wiley, 1956.

Estes, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.), Studies in Mathematical Learning Theory. Stanford: Stanford Univer. Press, 1959. Pp. 9-52.

Estes, W. K. and Suppes, P. Foundations of statistical learning theory, II. The stimulus sampling model. Tech. Rept. No. 26, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford Univer., Stanford, 1959.

Helseth, Inga Olla. The four fundamental arithmetical arithmetic processes in adults. Math. Teacher, 1922, 20, 265-273.

Knight, F. and Behrens, Minnie S. The Learning of the 100 Addition Combinations and the 100 Subtraction Combinations. New York: Longmans, Green, 1928.

Suppes, P. The psychological foundations of mathematics. Tech. Rept. No. 80, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford Univer., Stanford, 1965.

Suppes, P. and Atkinson, R. C. *Markov Learning Models for Multiperson Interactions*. Stanford: Stanford Univer. Press, 1960.

Suppes, P. and Groen, G. *Some counting models for first-grade performance data on simple addition facts*. Tech. Rept. No. 90, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford Univer., Stanford, 1966.

Footnotes

[1] To take care of the case when the observed p_i is either 0 or 1, we use the following transformation

$$Z_i = \begin{cases} \log (2n_i - 1) & \text{for } p_i = 0 \\ \log \dfrac{1}{2n_i - 1} & \text{for } p_i = 1 , \end{cases}$$

where n_i = the total number of subjects responding to item i. The exact form of this transformation is not important.

[2] The number of subjects or students shown in the various tables is always an approximation, with the exact number varying slightly from day to day.

[3] For reasons mentioned below, the first problem was deleted from each drill, leaving 19 problems per drill. The number of different daily drills in an analysis can be calculated by dividing the number of problems by 19.

[4] There are some repetitions of problems in Table 19, but all such repetitions occurred on different days.